CRANFIELD UNIVERSITY

JIE SHI

OBSTACLE DETECTION USING THERMAL IMAGING SENSORS FOR LARGE PASSENGER AIRPLANE

SCHOOL OF ENGINEERING
Digital Signal and Image Processing

MSc THESIS
Academic Year: 2011 - 2012

Supervisor: Al Savvaris
December 2012

CRANFIELD UNIVERSITY

SCHOOL OF ENGINEERING
Digital Signal and Image Processing

MSc Thesis

Academic Year 2011- 2013

JIE SHI

Obstacle Detection using Thermal Imaging Sensors for Large Passenger Airplane

Supervisor: Al Savvaris
December 2012

This thesis is submitted in partial fulfilment of the requirements for the degree of Master of Science

ABSTRACT

This thesis addresses the issue of ground collision in poor weather conditions. As bad weather is an adverse factor when airplanes are taxiing, an obstacle detection system based on thermal vision is proposed to enhance the awareness of pilots during taxiing in poor weather conditions. Two infrared cameras are employed to detect the objects and estimate the distance of the obstacle. The distance is computed by stereo vision technology. A warning will be given if the distance is less than the safe distance predefined. To make the system independent, the proposed system is an on-board system which does not rely on airports or other airplanes.

The type of obstacle is classified by the temperature of the object. Fuzzy logic is employed in the classification. Obstacles are classified into three main categories: aircraft, vehicle and people. Membership functions are built based on the temperature distribution of obstacles measured at the airport. In order to improve the accuracy of classification, a concept of using position information is proposed. Different types of obstacle are predefined according to different area at the airport. In the classification, obstacles are classified according to the types limited in that area.

Due to the limitation of the thermal infrared camera borrowed, images were captured first and then processed offline. Experiments were carried out to evaluate the detecting distance error and the performance of system in poor weather conditions. The classification of obstacle is simulated with real thermal images and pseudo position information at the airport. The results suggest that the stereo vision system developed in this research was able to detect the obstacle and estimate the distance. The classification method classified the obstacles to a certain extent. Therefore, the proposed system can improve safety of aircraft and enhance situational awareness of pilots.

The programming language of the system is Python 2.7. Computer graphic library OpenCV 2.3 is used in processing images. MATLAB is used in the simulation of obstacle classification.

Keywords:

Stereo Vision, Infrared Camera, Poor weather, Fuzzy logic

ACKNOWLEDGEMENTS

I would like to express my gratitude sincerely to those who have helped me during my research.

First of all, I would like to thank my company, Commercial Aircraft Corporation of China, Ltd (COMAC), who gives me the opportunity to study in Cranfield University for one year. I broadened my eyes and had a wonderful experience of my life this year.

I would like to thank Dr. Al Savvaris, my supervisor, who suggested me to take this project and provided enough support. Thank you for being a mentor and giving me enough autonomy.

My thanks also go to Mikael Mannberg, for your guidance and suggestion when I met difficulties during my research. It is always helpful to have a discussion with you.

Thanks to Dr. Lone Mudassir who was always willing to be of assistance.

The School of Engineering at Cranfield University has always supported my research. Thanks to Barry Walker and Ian Hakon, for making the plastic and metal pieces for camera calibration.

Thanks to Peter Anthony and STFC Rutherford Appleton Laboratory, for providing the instruments for me to develop the thermal stereo vision system.

Thanks to Stella Redman for the proof reading of my thesis and your hospitality. I wish all the best to you and your colleagues.

I would also like to thank my colleagues, COMAC students of 2012. It is a great honour to meet you and know each other in UK, thousands miles away from China. I will never forget the wonderful time we spent together.

Thanks to anyone I met in UK, whoever you are, wherever you come from, whatever you do.

Last but not least, I would like to thank my family for supporting me during my studies.

TABLE OF CONTENTS

ABSTRACT i
ACKNOWLEDGEMENTS iii
LIST OF FIGURES vii
LIST OF TABLES x
LIST OF ABBREVIATIONS xi
1 INTRODUCTION 1
1.1 Motivation 1
1.1.1 Influence of Poor Weather 1
1.1.2 Current System for Separation 1
1.1.3 The Proposed System 3
1.2 Aims and Objectives 4
1.3 Thesis Outline 4
2 LITERATURE REVIEW 7
2.1 Obstacle Detection System 7
2.1.1 Candidate Sensors 7
2.1.2 Sensor Fusion System 8
2.1.3 Stereo Vision System 9
2.2 Camera Models 10
2.3 Calibration Pattern 12
3 SYSTEM SPECIFICATION 15
3.1 Detection Range 15
3.2 FOV and IFOV 16
3.3 Camera Placement 17
3.4 User Interface 18
3.5 Programming Language 20
4 CAMERA MODEL AND CALIBRATION 21
4.1 Camera Model 21
4.2 Intrinsic Parameters 22
4.2.1 Camera Coordinate System 22
4.2.2 Principal-Point Offset and Image Skew 23
4.2.3 Lens Distortion 24
4.3 Extrinsic Parameters 25
4.3.1 Rotation Matrix 25
4.3.2 Translation Vector 26
4.4 Camera Calibration 27
4.4.1 Homography 27
4.4.2 Camera Calibration 28
5 STEREO VISION 29
5.1 Fundamental Theory 29
5.1.1 Coordinate Systems 29

5.1.2 Triangulation Principle 29
5.1.3 Epipolar Geometry 30
5.1.4 Essential and Fundamental Matrices 31
5.2 Stereo Calibration 31
5.3 Stereo Rectification 33
5.4 Stereo Correspondence and Re-projection 34
6 OBSTACLE DETECTION 37
6.1 Temperature Measurement 37
6.2 Image Processing 37
6.2.1 Image Normalization 37
6.2.2 Background Elimination 38
6.2.3 Depth Map Filtering 41
6.3 Obstacle Classification 42
6.3.1 Obstacle Temperature 42
6.3.2 Fuzzy Logic 43
6.3.3 Position Involved 45
6.4 Algorithm Flow 46
7 SYSTEM TESTING 49
7.1 System Setup 49
7.1.1 Hardware Setup 49
7.1.2 Baseline Distance 50
7.2 Stereo Calibration 51
7.2.1 Chessboard 51
7.2.2 Calibration Result 53
7.3 Testing Result 55
7.3.1 Distance Estimation Error 55
7.3.2 Imaging Under all Weathers 57
7.3.3 Temperature Range 59
7.3.4 Testing at the Airport 60
7.3.5 Simulation of Obstacle Classification 63
8 DISCUSSION AND CONCLUSION 67
8.1 Strengths and Limitations 67
8.2 Contributions 67
8.3 Suggestions of Future Work 68
8.4 Conclusion 69
REFERENCES 71
APPENDICES 75

LIST OF FIGURES

Figure 1-1: Distance between Aircrafts on Taxi-way 2
Figure 1-2: Radar and User Interface of SMR System 2
Figure 1-3: Principle of ADS-B System 2
Figure 1-4: User Interface of A-SMGCS System 3
Figure 2-1: Pinhole Model 11
Figure 2-2: Thin Lens Model 11
Figure 2-3: Thick Lens Model 11
Figure 2-4: Full Lens Model 11
Figure 2-5: Printed Chessboard Heated by Infrared Radiation 12
Figure 2-6: Black Dots Printed on Foam Core Board 12
Figure 2-7: Mask Based Calibration Board 13
Figure 2-8: Black Metal Frame with White Plastic Inserts 13
Figure 2-9: Calibration Board with Miniature Bulbs 14
Figure 2-10: A Chessboard with a Grid of Wires 14
Figure 3-1: Minimum Distance of Separation 15
Figure 3-2: FOV and IFOV 16
Figure 3-3: Sensor Spatial Resolution (10x10 pixels as detectable area) 17
Figure 3-4: Camera Placement 18
Figure 3-5: User Interface 19
Figure 3-6: HUD in Aircraft 19
Figure 4-1: Pinhole Camera Model 21
Figure 4-2: Pinhole Camera Model (Image panel shifted) 21
Figure 4-3: Camera Coordinate System 23
Figure 4-4: Principle Point offset and Image Skew 23
Figure 4-5: Lens Radial Distortion 24
Figure 4-6: Lens Tangential Distortion 24
Figure 4-7: Rotate around One Axis 25
Figure 4-8: Translation of Coordinate System 26

Figure 4-9: Rotation and Translation 26
Figure 4-10: Homography..... 27
Figure 4-11: Images Captured in the Calibration..... 28
Figure 5-1: Stereo Vision Coordinate System 29
Figure 5-2: Triangulation Principle..... 29
Figure 5-3: Epipolar Geometry 30
Figure 5-4: Images Captured by Two Cameras..... 32
Figure 5-5: Align the Two Cameras into One Viewing Plane Mathematically.. 33
Figure 5-6: Transform the Left and Right Images in One Plane 33
Figure 5-7: Principle of Rectification..... 34
Figure 5-8: Correspondence..... 34
Figure 6-1: Temperature Measurement of Infrared Camera..... 37
Figure 6-2: Infared Image with Different Temperature Scale..... 38
Figure 6-3: Background Elimination 39
Figure 6-4: The Influence of Background Elimination..... 40
Figure 6-5: Depth map filtered based on contour area..... 42
Figure 6-6: Membership Functions..... 44
Figure 6-7: Measured Temperature Distribution of Obstacles..... 45
Figure 6-8: Algorithm Flow 46
Figure 6-9: Image Processing Time 47
Figure 6-10: Image Processing Time (without background elimination) 47
Figure 7-1: Stereo Vision System with a Pair of Infrared Camera 49
Figure 7-2: Baseline Distance 50
Figure 7-3: Depth and Disparity..... 51
Figure 7-4: The Common Field of Two Cameras 52
Figure 7-5: Thermal Image of Heated Chessboard 52
Figure 7-6: Two kinds of Chessboard..... 53
Figure 7-7: Thermal Infrared Camera Calibration..... 53
Figure 7-8: Thermal Images Captured in Stereo Calibration 54

Figure 7-9: Distance Estimation Error Testing 56
Figure 7-10: Actual Distance and Measured Distance 57
Figure 7-11: Error of Distance Estimation 57
Figure 7-12: People in Fog 58
Figure 7-13: Airplanes in Fog 58
Figure 7-14: Airport at Night 58
Figure 7-15: Airport in Rain Day 59
Figure 7-16: Airport on Sunny Day 59
Figure 7-17: Human body's Temperature (Cloudy) 60
Figure 7-18: Human body's Temperature (Fog) 60
Figure 7-19: Obstacle Detecting and Depth Map Filtering 61
Figure 7-20: Moving Target Detecting 62
Figure 7-21: Position Distribution of Objects 63
Figure 7-22: Map of Airport 64
Figure 7-23: Concept of Obstacle Classification 64
Figure 7-24: Temperature Distribution of Objects 65

LIST OF TABLES

Table 6-1: Temperature range of obstacles 43
Table 6-2: Membership function Settings 44
Table 7-1: Computer Specification 50
Table 7-2: Intrinsic Calibration Results 55
Table 7-3: Relative Extrinsic Calibration Results 55
Table 7-4: Performance Comparisons 66

LIST OF ABBREVIATIONS

ADS-B	Automatic dependent surveillance-broadcast
A-SMGCS	Advanced Surface Movement Guidance and Control System
ATC	Air Traffic Control
BB	Bounding Box
COMAC	Commercial Aircraft Corporation of China, Ltd
EO	Electrical Optical
GPS	Global Positioning System
HUD	Heads-Up Display
RGB	Red Green Blue
SMR	Surface Movement Radar
IR	Infrared Radiation
MMW	Millimetre-wave
YCbCr	Luma component, blue-difference Chroma component, red-difference Chroma component

1 INTRODUCTION

'It doesn't matter what the conditions are, it's still invaluable to see out of the window' [1]

David Learmount
Flight International

1.1 Motivation

1.1.1 Influence of Poor Weather

Poor weather such as fog is an adverse factor for passenger airplanes and airports. For airplanes, pilots' situational awareness is weakened by low visibility. The low visibility makes it hard for pilots to detect obstacles and estimate distance between airplanes. Pilots have to taxi slower and pay more attention to the situation, which might slow the whole process down at the airport.

Moreover, since pilots are less able to see the potential dangers, the traffic controller has to increase the space between airplanes. The purpose is to give the pilot more time to react if anything goes wrong. However, the result is fewer airplanes can use the airport and the capacity of the airport might decrease. The decrease of capacity will cause delay and cancellation of flights, thus causing chaos at airports, resulting in unhappy passengers. In the run-up to Christmas in December 2006, thousands of travellers had to cancel their holiday due to a dense fog which caused disruption to flights.

Collisions are more likely to happen with low visibility in poor weather conditions. In October 2001, a Scandinavian SAS passenger jet collided with an airplane which was crossing its path in fog at Milan's Linate airport. 118 people in both airplanes were killed. Although the collision occurred due to a combination of factors, it would have been helpful if the pilot could have seen what was happening in front of the airplane. Then, immediate actions could have been taken and accident could have been avoided.

1.1.2 Current System for Separation

The distance between aircrafts increases when taxiing at the airport in poor weather conditions. Figure 1-1 shows the separation of aircraft when taxiing at the Heathrow Airport. In poor weather conditions, the distance is increased due to the consideration of safety. Due to the low visibility in poor weather, pilots and traffic controllers have to rely more on systems than their eyes. Current systems such as Surface Movement Radar (SMR), Automatic Dependent Surveillance-Broadcast (ADS-B) or Advanced Surface Movement Guidance and Control System (A-SMGCS) can be used for separation in poor weather conditions.

SMR is mainly for detecting airplane and vehicles on the surface of an airport. Figure 1-2 shows the radar and user interface of SMR system. It can be used under low visibility conditions such as night since it uses radio waves as signal.

It interprets the reflected signals from targets and estimates the distance. The movement of airplane and vehicles can be monitored by air traffic controllers to supplement visual observations. The user interface always employs a map of the airport. Typically, SMR presents as a video blip overlaid onto a plan view map of the airport. Run-way and taxi-way and buildings are also shown on the map. Targets can be identified and a warning will be provided in the event of potential conflicts between airplanes.

Figure 1-1: Distance between Aircrafts on Taxi-way (image source: [2])

(a) Radar of SMR **(b)** User Interface of SMR

Figure 1-2: Radar and User Interface of SMR System (image source: [3][4])

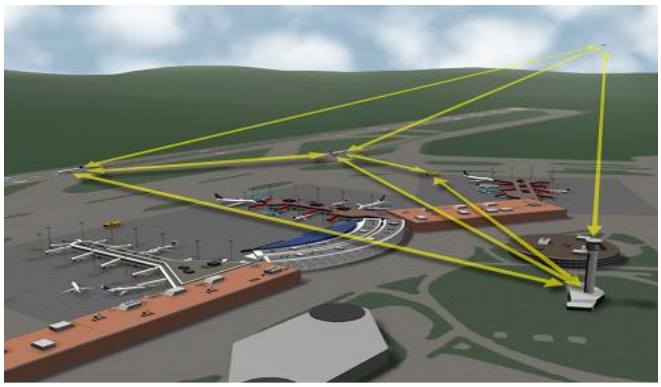


Figure 1-3: Principle of ADS-B System (image source: [5])

Figure 1-3 shows the principle of ADS-B system. There are two fundamental components of ADS-B system. One is the GPS navigation source and the other is the broadcast communication link. Both equipment need to be installed in the aircraft. First, the aircraft uses a GPS signal receiver to get its position. Then, the position information will be integrated with other information, such as altitude and speed. The combined information will be broadcasted to other aircraft and ATC centres in real time. Thus, ADS-B enhances safety by broadcasting the aircraft position to ATC and other aircraft. It can be used for ground surveillance in poor weather conditions.

Figure 1-4: User Interface of A-SMGCS System (image source: [6])

A-SMGCS is a multifunctional system which can be used to provide routing, guidance and surveillance for the control of aircrafts and vehicles. Figure 1-4 shows the user interface of A-SMGCS system. Information from SMR, ADS-B, GPS and ground sensors can be integrated into the system and sent to other aircraft and ATC centres. It allows for enhanced low visibility operations within the aerodrome visibility operational level. A380 has been equipped with an A-SMGCS system to improve the safety of surface movement. A moving map displays the location of all the aircraft and vehicles at the airport. Even if the visibility is low in poor weather conditions, the pilot will able to identify the location of other airplanes displayed on the screen.

1.1.3 The Proposed System

The systems described in Chapter 1.1.2 can be used for separation in poor weather conditions. However, they are not completely independent systems. As illustrated, SMR is mainly used by the traffic controller to supplement visual observation. Pilots need to wait for the instructions sent by the traffic controller. Time might be wasted during the communication between pilot and controller. In poor weather, the increased work load for the traffic controller may result in mistakes being made. As for ADS-B and A-SMGCS, the location of other airplanes comes from the information received from other airplanes or ATC, which makes the system rely on facilities at the airport and other airplanes. Compared with the coordinates given by these systems, vivid images are better in enhancing pilots' awareness.

Thus, a vision based obstacle detection system is proposed in this thesis. Obviously, vision based systems are more like human eyes. Image can provide

more information than coordinates from radar. Once pilots see the obstacle, actions can be taken immediately instead of waiting to be told by the traffic controller. The system should provide the distance of obstacles and a warning should be given if the distance is less than the safe distance predefined. Also, the type of obstacle should be identified and displayed by the system. Last but not least, the system should be an on-board and independent system which does not rely on airports or other airplanes.

1.2 Aims and Objectives

The aim of this project is to develop an obstacle detection system to enhance pilots' situation awareness using externally mounted sensors that could operate in poor weather conditions. The main objectives of the projects are:

1. Define system performance specification
2. Select and implement image processing algorithms
3. Implement obstacle detection and classification algorithms
4. Test and evaluate the system performance

1.3 Thesis Outline

The following chapters illustrate the methods adopted to achieve the aims of this research.

The first part in Chapter 2 reviews the obstacle detection system was developed using vision sensors. Relative sensors and vision systems that could be used for this research is discussed. Computer vision using thermal cameras is selected as preferred technology. Then, the cameral models used in the stereo vision system are studied. A simple pinhole model is selected as the camera model. Since thermal cameras are only sensitive to heat, the calibration pattern is different from optical cameras. The final section reviews the calibration pattern developed for thermal cameras.

Chapter 3 defines the system performance specification. In first part, the detecting distance is defined based on the safe distance between aircrafts, which allows enough time for the airplane to stop. Then the influence of lens and size of detector for camera imaging is analysed. A lens with narrow angle of view and a detector with high resolution are preferred. The third part discusses the placement of cameras. Angle of view, vibration and detecting range are considered in the discussion. Last part presents a brief introduction of user interface and programming language of the system.

Chapter 4 illustrates the camera model and calibration. The first part presents the pinhole camera model and perspective projection. Then, parameters of camera including intrinsic parameters and extrinsic parameters are introduced. The last part presents the homography theory in calibration.

Chapter 5 focuses on the stereo vision theory. The coordinate system of stereo vision system is introduced in the first part. Then, the basic theory of stereo

vision is introduced, including triangulation principle, epipolar geometry, essential and fundamental matrices. Then, a brief introduction the calibration of stereo vision system is presented. The last part illustrated the rectification of stereo images, correspondence and calculation of depth map.

Chapter 6 address the obstacle detection of the system. The temperature measurement of thermal cameras is introduced in the first part. Then, the second part introduces the implement of image processing and filtering algorithms for thermal images. Obstacle classification method developed for the system is introduced in the third part. Fuzzy logic is employed in the obstacle classification. A concept of position information involved classification is also addressed. The last part highlights the algorithm flow of the system.

Chapter 7 presents the system testing and testing result of the proposed system. Part one introduces the system setup. The selection of baseline distance is discussed in this part. The selection takes detecting range and distance resolution into consideration. The second part discusses the stereo calibration in real situation. A big chessboard is made due to the increased baseline distance of two cameras. Also, a new chessboard made of plastic and aluminium pieces is made to obtain better image quality.

Chapter 8 discusses the result of testing, pointing out the strengths and limitations of the system. The contribution is highlighted in the second part. Suggestions for future research are given in part three. Finally, the main conclusions of this work are discussed.

2 LITERATURE REVIEW

2.1 Obstacle Detection System

2.1.1 Candidate Sensors

A large number of obstacle detection systems based on vision sensors have been developed in the past years. Sensors used in vision system are mainly electro-optical (EO) camera and thermal infrared radiation (IR) camera. In order to detect the distance, millimetre-wave (MMW) radar and laser scanner are commonly employed in these systems.

Electro-optical cameras are widely used in surveillance and obstacle detection systems. The low cost make them ideal sensors in different applications. Images provided by electro-optical cameras are replicate real life images. With proper lens, electro-optical cameras can see much further than human eyes. Obstacles can be identified by pilots directly or by the system using obstacle detection algorithms. However, the change of illumination and low visibility conditions will affect its performance. Additionally, strong light beams from other airplanes or vehicles can also influence the imaging quality of electro-optical sensors.

Thermal infrared cameras are commonly used in low visibility conditions. They have been used in wildfire detection [7], negative obstacle detection [8] and people tracking for mobile robots [9]. There are two types of sensors employed in thermal imaging cameras. One type is the uncooled sensors working in the long-wave infrared (LWIR). Another type of sensor is cooled detectors. The cooled sensors are sensitive to small temperature differences. The performance of thermal infrared cameras in poor weather conditions is discussed in [10]. Fog is classified into four categories according to visual range. The detection range is calculated by an atmosphere model. The simulation results show that generally the detection range of IR camera is better than that of visual camera.

Millimetre-wave (MMW) radar can be used in distance measurement. It can penetrate poor weathers (such as fog and dust) and measure distance of obstacles. There are two kinds of MMW radars, passive MMW radar and active MMW radar. The passive MMW radar detects the inherent electromagnetic radiation of objects. The active MMW radar needs to emit MMW and then measure the reflected signal. High reflectivity objects such as metallic objects are ideal for detecting. Objects position and geometry can also be measured by MMW radar. However, the image quality such as contrast and resolution depends on several factors such as beam width and antenna aperture. Generally, MMW radar has low directional resolution which will influence the measurement accuracy.

Laser scanners use a controlled laser beam to measure the distance pointing in every direction. The distance can be measured to form a depth map. It is used to rapidly capture shapes of objects, buildings and landscapes. A popular product is Kinect developed by Microsoft. An infrared laser projector and a monochrome CMOS sensor is employed in this device. The infrared laser

makes Kinect capture data under low visibility conditions. However, similar to the MMW radar, the laser scanner provides only an incomplete representation of the scene.

From the above, it can be seen that the sensors discussed have advantages and disadvantages. In order to have high spatial resolution and better weather penetration capability for the proposed system, a simple way is to use sensor fusion technology.

2.1.2 Sensor Fusion System

The purpose of sensor fusion system is to make use of the advantages of different sensors. The combination can improve system performance and system reliability. For instance, a combination of an electro-optical camera and millimetre-wave radar can make the system have both excellent spatial resolution and distance measuring ability. A proper fusion of sensors can make the system suitable for the application of this research.

One kind of sensor fusion is millimetre-wave radar and an electro-optical camera. In [11], the proposed system use sensor fusion of a video camera and a millimetre-wave radar. The radar is used to measure the distance while the boundaries of road are extracted from image sequences. Then, the images captured by the camera are enhanced by the distance information measured by radar. In order to improve the accuracy, data acquired by radar can be divided to clusters, segmented and visualized on the images captured by the camera [12]. However, the calibration between radar and camera is difficult to carry out. A fusion-tracker and pedestrian classifier is proposed in [13]. Data captured from EO and thermal infrared cameras are used to build a multi-modal distribution of colour and temperatures. In [14], the researcher proposed an obstacle detection system based on three sensors for civil helicopters. Three kinds of sensors, an EO camera, an IR camera and MMW radar are combined for the system. Data gathered from three sensors are fused into one image highlight the power line. In [15], the fusion between stereo vision and a laser scanner is employed in a road obstacle detection system. The laser sensor is used to detect and track the obstacles and the stereo vision is employed to confirm what the laser tracks. Comparing with using a single camera, the stereo vision can provide more information.

For the application of obstacle detection in poor weather and distance measurement; electro-optical cameras suffer in low visibility conditions. Thus, the sensor fusion of thermal infrared camera and MMW radar, a stereo vision system based on two thermal infrared cameras, or fusion of stereo vision and MMW radar can be used to detect the obstacle under poor visibility conditions. Objects in thermal image can be identified by pilots and the distance of obstacle can be measured by the MMW radar. Sensor fusion of the EO camera, IR camera and MMW radar is also suitable for the application. However, the cost of the system needs to be taken into consideration. When there are more sensors, the cost of the whole system will rise significantly. Moreover, the calibration of locations between different sensors is an issue. Since the distance

can be calculate by stereo vision technology, a stereo vision system is another solution for this research.

2.1.3 Stereo Vision System

Stereo vision has been widely studied and applied in road obstacle detection [16], vehicle detection [17] and airports runway detecting[18]. Most of the stereo vision uses a pair of electro-optical camera as sensors. In [18], the stereo vision is used to detect the run-way for guidance and control of aircraft landing during an approach. The edge line of the run-way is detected by Hough transform firstly. Then, distance and touchdown point is estimated. In [19], an obstacle detection system based on stereo vision is proposed to avoid wingtip collision in aerodrome Areas. In this system, a pair of electro-optical camera is mounted on the wingtips to track the obstacles and estimate the distance between obstacle and wingtip. The test results show the system is able to detect and track generic obstacles. One of the limitations of the system is that the results obtained are not always predictable in low visibility conditions. The performance of system deteriorates due to the increasing image noise when the system is in low visibility condition.

In [20], stereo vision system using thermal infrared cameras was developed for unmanned ground vehicle perception. A pair of high-quality thermal infrared is used in stereo ranging and obstacle detection. Three main factors that influence the quality of stereo range data are analysed. One issue of stereo ranging is the motion of thermal cameras will reduce the image quality, which may cause the loss of stereo data. High pitch vibration will cause the image to blur and high speed will cause low texture. Since the calibration of thermal cameras is different from optical camera, poor calibration will also cause stereo range data lose.

Although two cameras are adequate in estimating distance from image data, detection accuracy can be improved by using more than two cameras. In [21], the advantage of using more than two cameras is illustrated. Three cameras in an “L” configuration are employed to detect small obstacles in long range. The object which is about 14cm high could be detected at range of $100\,m$. In [22], a concept of three thermal infrared cameras based obstacle detection system is proposed. The purpose of the system is to detect small obstacles in the pantograph gauge. Since the performance of electro optical camera degrades when daylight is not sufficient, thermal infrared cameras are employed to detect the obstacle at night and in tunnels. The analysis shows that the proposed system is able to detect small obstacles in the pantograph gauge under low visibility situations by using three thermal infrared cameras.

In this research, thermal infrared cameras are considered as the sensors. Since thermal infrared camera is only sensitive to heat, obstacles such as aircrafts can be easily detected according to their temperature. With the stereo vision technology, the distance information can be calculated. As for the number of cameras, the detecting distance of two EO cameras is about 50 metres [16] while 100 metres can be achieved by three cameras [21]. In this research it is

not necessary to detect small obstacles with three cameras. Thus, two thermal infrared cameras can meet the system's requirements.

2.2 Camera Models

In computer graphics, several camera models are commonly employed to approximate physical optical systems [23]. They are pinhole model, thin lens model, thick lens model and full lens model. These models usually approximate a system of parallel lenses.

Pinhole model is a simple camera model. In this model, all the light rays from the scene pass through a single point, shown in Figure 2-1. The image panel shows the image of objects captured by pinhole model. It can be seen that each point on the object can only form one light ray which can pass through the pinhole. The result is the image of object always appears in focus. In reality, since only a little part of light goes through the pinhole, the image will be too dim to observe. Thus, lenses are employed to allow more light go through the pinhole.

The thin lens model assumes the lens have an infinitesimal thickness. It is an ideal model that the thickness can be ignored even if the lenses are convex or concave. In Figure 2-2, the optical axis goes through the centre of the lens. The focal point F located at optical axis of the lens and the plane at the lens is the principal plane. Any light ray traveling parallel to the axis will proceed toward the focal point F' after refraction. The focal length f is the distance between the centre of the lens and a focal point F. The relationship between f and l can be written as

$$\frac{1}{l}+\frac{1}{l'}=\frac{1}{f} \qquad \textbf{(2-1)}$$

where l is the distance between object and lens, l' is the distance between lens and image. In the thin lens model, it is assumed that the thickness of the lens can be ignored since it is relatively small compare with the focal length. In some cases such as the photographic system, the lens cannot be simplified as a thin lens. However, sometimes it is possible to simplify these lenses as ideal thick lenses.

Figure 2-3 shows a typical thick lens. It is a single lens with two spherical surfaces. The two surfaces are separated by an appreciable distance. There are two principal planes H and H'. F and F' are the corresponding focal points. The dashed lines show the actual paths of rays passing through the lens. The exit directions for the rays can be calculated similar to the way of the thin lens.

Since the thin lens and thick lens models are not sufficient for describing the geometry and radiometry, a physically-based camera model that simulated a system of lenses is described in [24]. Figure 2-3 shows the typical full lens model. It can be seen that more elements are employed in this lens system, including simple lenses, stops and aperture. With this model, the geometry of

image formation, the change of geometry that occurs during focusing, exposure levels can be taken into account.

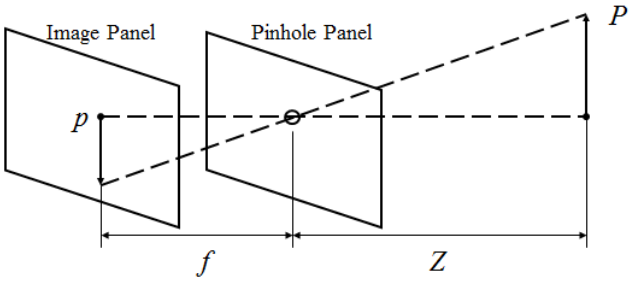


Figure 2-1: Pinhole Model (image source: [24])

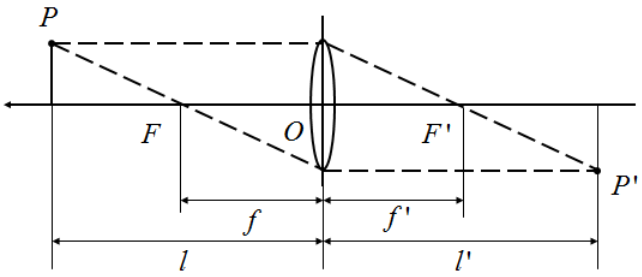


Figure 2-2: Thin Lens Model (image source: [24])

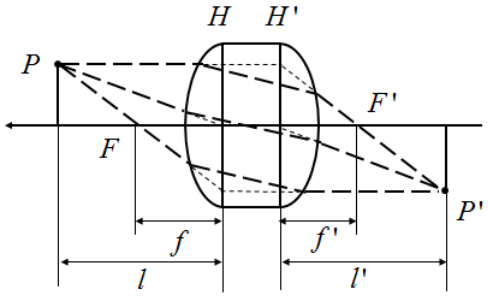


Figure 2-3: Thick Lens Model (image source: [24])

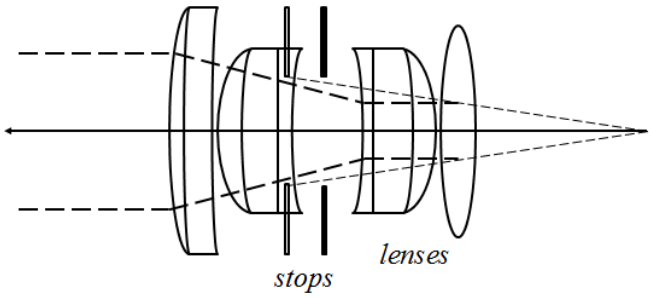


Figure 2-4: Full Lens Model (image source: [24])

2.3 Calibration Pattern

Calibration patterns are used in the camera calibration process in stereo vision. The parameters of the camera can be calculated by capturing the calibration pattern from different orientations. A chessboard is the most common pattern for optical camera calibration. The chessboard used can be of any size N by M. The black-white intersections of the chessboard squares can be easily recognized. Since thermal infrared camera is only sensitive to heat, contrast between white and black pieces of the chessboard is not obvious in thermal image. If the thermal image of chessboard has the same contrast as visual image, the rest work of calibration is the same as optical camera. Different kinds of calibration pattern have been made by researchers to provide high contrast in thermal images. These calibration patterns can be divided into three types.

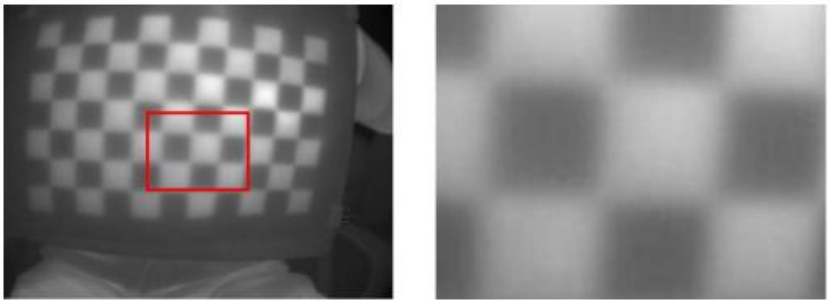


Figure 2-5: Printed Chessboard Heated by Infrared Radiation (image source: [25])

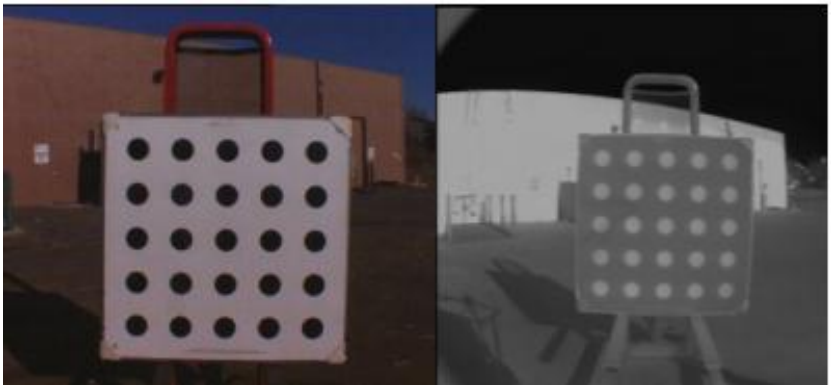


Figure 2-6: Black Dots Printed on Foam Core Board (image source: [20])

The first type is the printed calibration pattern heated by radiations (such as sun or infrared radiation). In [26], the researcher proposed a method to calibrate the thermal infrared camera by using a printed chessboard and an infrared radiation. Due to white and black colour have different absorb rate of IR light, the heated chessboard can make the calibration similar to conventional calibration. Figure 2-5 shows the chessboard heated by an infrared radiation. The sun is a perfect radiation while the IR lamp has the limitation of field of operation. Figure 2-6 shows another calibration pattern. By heating a foam core board printed with black dots under the sun, thermal contrast between dots and background will occur. However, this kind of pattern works only when there is a radiation source

which can heat the pattern uniformly. Also, it can be seen that the conductivity of the heat will make the edge less sharp. It will affect the detection precise of features within the pattern.

The second type always employs a mask to cover part of the infrared rays from the radiation inside the pattern. Bulbs are always used as radiation sources. A geometric mask is proposed as an alternative calibration pattern in [25]. High thermal contrast can be provided, which is shown in Figure 2-7. A clustering-based algorithm is employed to locate the calibration features of the board. A mean re-projection error up to 78% lower can be achieved by using this kind of chessboard. Figure 2-8 shows the calibration pattern made of black metal frame with white plastic inserts. On the back of plastic inserts, there are miniature bulbs attached. It can be seen in the right image, the calibration pattern is not heated by bulbs uniformly. The advantage of this pattern is it does not rely on an outside radiation source like sun and the blurred edge produced by conductivity of heat can be eliminated.

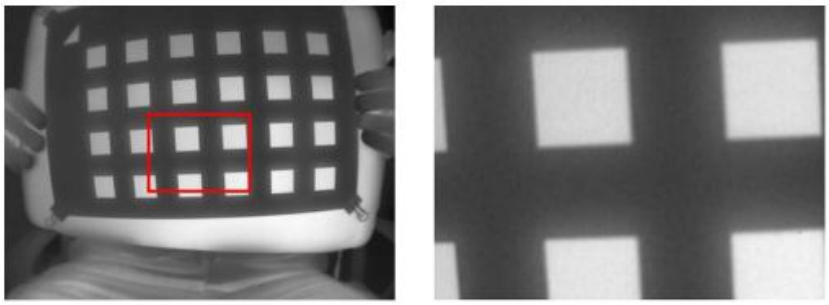


Figure 2-7: Mask Based Calibration Board (image source: [27])

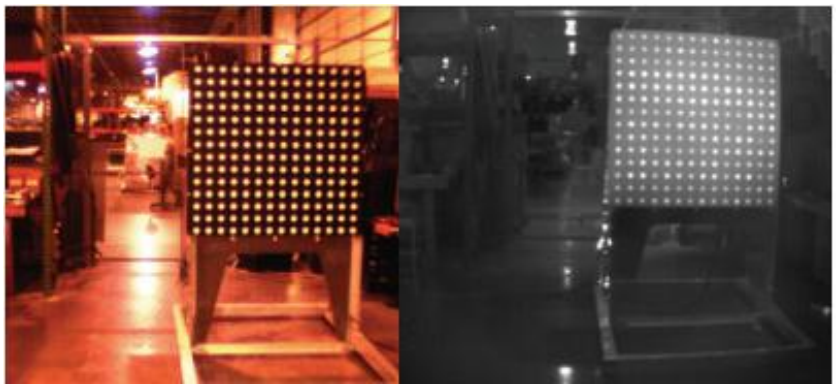


Figure 2-8: Black Metal Frame with White Plastic Inserts (image source: [20])

The third type is the structural light. Figure 2-9 shows the calibration board employed in [27]. A board with 25 holes with bulbs inside is designed to calibrate the IR camera. Heat and light can pass through the holes when the power is turned on. The IR cameras can capture the heat coming from the holes on the board. Figure 2-10 shows a calibration pattern that can be used for optical and thermal infrared cameras. A thin transparent sheet printed with a chessboard pattern is placed on an aluminium panel. A grid of wires is put on the back of the chessboard. The wires cross at the corners of the chessboard.

The printed chessboard is for the optical camera while the wires can be detected by thermal infrared cameras when current passing by. Optical camera and thermal camera can be aligned precisely by this structure.

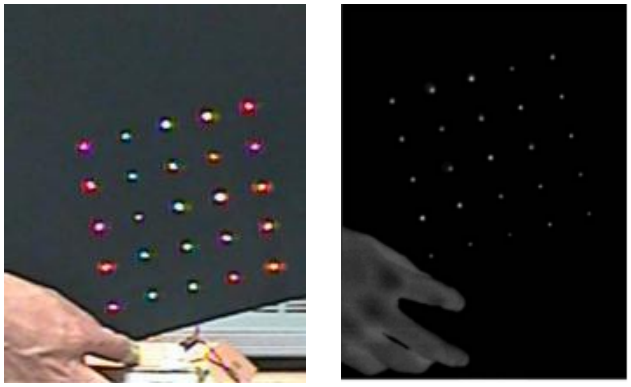


Figure 2-9: Calibration Board with Miniature Bulbs (image source: [27])

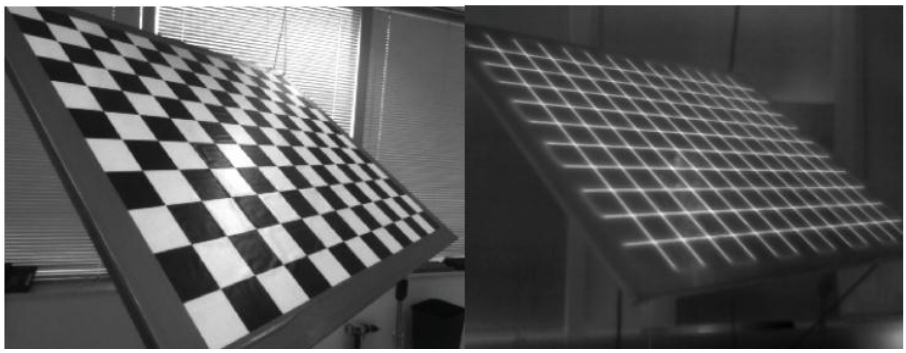


Figure 2-10: A Chessboard with a Grid of Wires (image source: [20])

Since a chessboard is used as the calibration pattern for optical image in OpenCV, it is easier to use a chessboard. The mask-based pattern and the board with bulb need extra algorithm to locate the calibration point. The algorithm in OpenCV can be applied in thermal imaging directly. Thus, a chessboard is employed in thermal infrared camera calibration in this thesis.

3 SYSTEM SPECIFICATION

3.1 Detection Range

The calculation of detection range is based on the airplane taxiing speed, system processing time, pilots' reaction time and braking distance. The taxiing speed varies according to different conditions [28]. In busy areas, the taxiing speed should no more than walking speed of human. The average walking speed is of human is about 1.4 m/s^2 [29][30]. In open areas, the taxiing speed may be increased. Typically, the taxiing speeds of airplane are from 5 to 20 knots [31] (2.57~10.28 m/s^2). System processing time includes camera frame capture time and image processing time. Assuming the capturing rate of camera is 15 Hz, in order to make the system real time, the whole processing time of the system should be less than 1/15 s. The typical pilot reaction time is about 1-2 s [32]. The mean deceleration in an overweight A380-800 braking test is 3.62 m/s^2 [33].

Considering a worst case scenario, a maximum speed of 20 knots (10.28 m/s^2) on taxi-way, pilot reaction time of 2 s and deceleration rate 3.62 m/s^2 are adopted in this thesis. Then, the minimum distance can be calculated by the following equation:

$$d_{stop} = vt_{all} + d_{brake} \tag{3-1}$$

$$d_{stop} = vt_{all} + \frac{v_i^2 - v_f^2}{2a_{brake}} \tag{3-2}$$

where d_{brake} is the braking distance, v_f is the final speed (0 m/s), v_i is the initial speed (10.28 m/s), t_{all} is the overall time (including capturing, image processing and typical pilot reaction), a_{brake} is the average braking deceleration rate.

The result calculated shows the minimum distance between two airplanes is 35.84 m. The detecting range should take distance between the cockpit and the system into account when the camera is not placed in the cockpit of the airplane. Thus, the working distance of the system is defined as 40~80 m in this research.

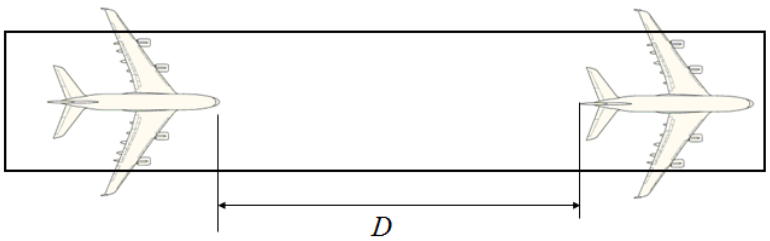


Figure 3-1: Minimum Distance of Separation

3.2 FOV and IFOV

The basic performance of a thermal infrared camera is how small an object or area can be detected and accurately measured at a given distance. This can be described by two parameters, Field of View (FOV) and Instantaneous Field of View (IFOV).

FOV is the size of area captured in the scene, shown in Figure 3-2 (a). It can be seen wh plane is the image plane while WH plane is the field of view. For a given distance D and camera lens ($\alpha° \times \beta°$), the FOV of camera can be calculated by

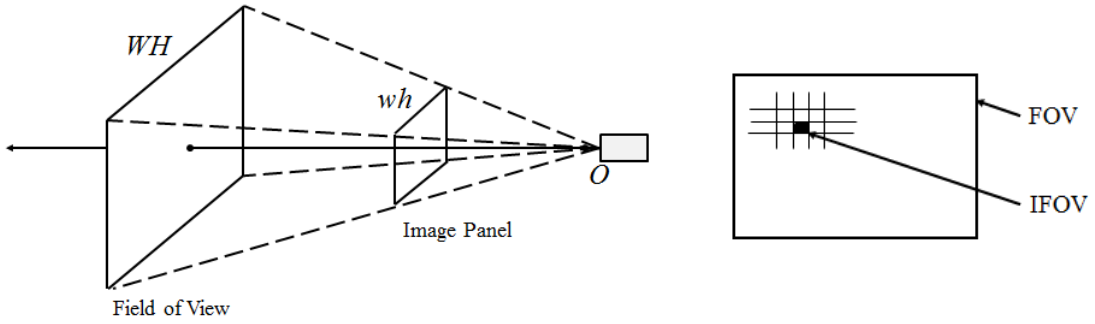
$$W = 2D\tan(\alpha/2) \quad \textbf{(3-3)}$$

$$H = 2D\tan(\beta/2) \quad \textbf{(3-4)}$$

where W and H are the width and Height of WH plane.

IFOV is used to describe the spatial resolution of a camera's detector. For a given distance, it determined the smallest detectable object. The IFOV can be calculated by the field of view and the size of detector. Figure 3-2 (b) shows the concept of IFOV. Assumin the size of image is $w \times h$ pixels, given distance D and camera lens ($\alpha° \times \beta°$), the smallest detectable area in WH plane is

$$s = 4D^2 \tan(\alpha/2)\tan(\beta/2)/wh \quad \textbf{(3-5)}$$



(a) FOV **(b)** IFOV

Figure 3-2: FOV and IFOV (image source: [34])

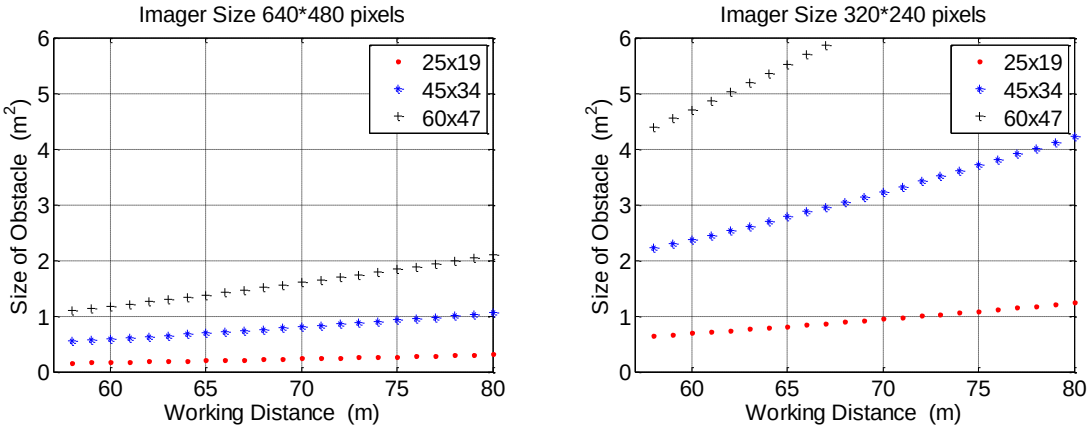
Assuming $10 \times 10 pixels$ is the smallest area in the image that can be identified by eyes, and then the minimum size of object can be figured out at a given distance. The relationship between the minimum size of obstacle in WH plane is calculated and shown in Figure 3-3. Three typical camera lens (angle of view $25^o \times 19^o$, $45^o \times 34^o$ and $60^o \times 47^o$) and two kinds of resolution ($640 \times 480 pixels$ and $320 \times 240 pixels$) are compared.

It can be seen in Figure 3-3 (a), for a camera with a narrow angle of view ($25^o \times 19^o$), the detectable area is about $0.25m^2$ at the distance of $80m$. For a camera with wide angle of view ($60^o \times 47^o$), the detectable area is about $2.1m^2$

at the distance of $80m$. Thus, with a same resolution, a camera with a narrow angle of view has a higher spatial resolution than camera with wide angle of view

It also can be seen from Figure 3-3, for a high resolution camera (in Figure (a)), the detectable area is about $0.25m^2$ at the distance of $80m$ (sensor size $640\times480\,pixels$, angle of view $25^o\times19^o$). For a low resolution camera (in Figure (b)), the detectable area is $1.25m^2$ at the distance of $80m$ (sensor size $320\times240\,pixels$, angle of view $25^o\times19^o$). Thus, with the same lens, the low resolution camera has a lower spatial resolution than the high resolution camera.

Thus, a camera with narrow angle of view and high resolution is ideal for detection. In the civil market, the resolution of FLIR SC640 is $640\times480\,pixels$. The disadvantage of a narrow angle of view is an obstacle may fall outside view of the camera lens.



(a) Sensor size $640\times480\,pixels$ **(b)** Sensor size $320\times240\,pixels$

Figure 3-3: Sensor Spatial Resolution (10x10 pixels as detectable area)

The unit of WH plane is m while the unit of wh plane is $pixel$. A simple relationship between the angle of view ($\alpha°\times\beta°$), length (f) and detector size ($w\times h$) can be written as:

$$f=\frac{(1/2)x}{\tan(\alpha/2)} \qquad \textbf{(3-6)}$$

Assuming the camera has an angle of view $25^o\times19^o$, detector size $8mm\times6mm$ and sensor resolution $640\times480\,pixels$, then the focal length is $1443\,pixels$.

3.3 Camera Placement

Three positions are considered in mounting the camera. These positions are the tailfin, wingtip and cockpit.

Based on the assumption that detection distance is not a problem, the advantage of placing the camera in the tailfin is that a larger area of the airplane can be covered. Even if the angle of view of camera is narrow, there is still a large area that can be detected in front of the airplane. It is even possible to monitor the wingtips. However, it has been calculated that the detection range is an issue when the working distance is long. The distance error will increase dramatically when the detecting distance increases.

Mounting the camera on the wingtip is another choice. The working distance is not too long when the length of front fuselage and minimum distance are added. A serious issue of placing the cameras in this location is the vibration of the wing. As discussed in the literature review, one of the factors which cause loss of stereo range data is the image blur due to high pitch rate. The vibration will affect the image quality and then result in wrong distance estimation. Although an image stabilizing filter can be used to reduce the influence, the vibration can disturb the synchronization of image capturing on both cameras.

The advantage of placing the camera in the cockpit is that working distance can be minimized. Since the cockpit is the nearest position to the obstacle, there is no need to consider the length of aircraft before the mounting location. Since the distance resolution decreases dramatically when the detecting distance increased, having a short detecting distance will benefit the spatial resolution. By placing two pairs of cameras of different orientation, the system can have a wider view.

Thus, the cockpit is the preferred place for mounting the stereo vision system.

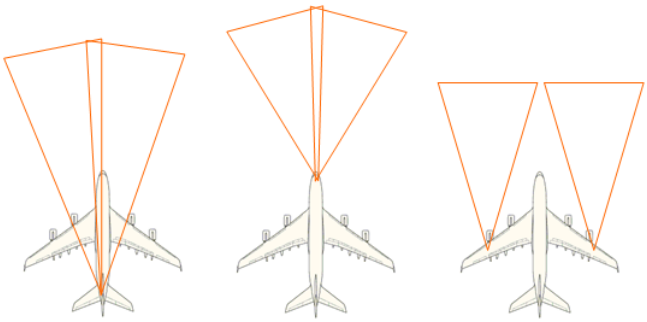


Figure 3-4: Camera Placement

3.4 User Interface

The system should provide a friendly user interface for the pilots. Figure 3-5 shows the user interface of the proposed system. Obstacles are highlighted by a bounding box (BB) in the thermal image. The bounding box is a bright-colour rectangle. The size of rectangle is based on contour area of obstacles. The colour of rectangle will change according to the distance estimated. There are three colours for the BB, which are green, yellow and red. Green means the obstacle distance is more than 60 meters. If the distance is within 40-60 meters, the BB will be drawn in yellow. As long as the distance is less than 40 meters,

the bounding box will be drawn in red and a warning message will pop up and flash.

Also, the obstacle type, distance and temperature are displayed on the right of the bounding box. Letter ‘P’ stands for people, letter ‘V’ stands for vehicle and letter ‘A’ stands for aircraft. The left-top corner of the image shows the taxiing speed while the right-top corner shows the frame rate. The left bar shows the speed of aircraft and speed limitation range. The right bar shows the distance of obstacle and a distance range that will trigger the warning message.

The user interface will be displayed on the Heads-Up Display (HUD). Figure 3-6 shows a typical HUD. A thermal image will be displayed on the HUD to enhance pilots’ situational awareness.

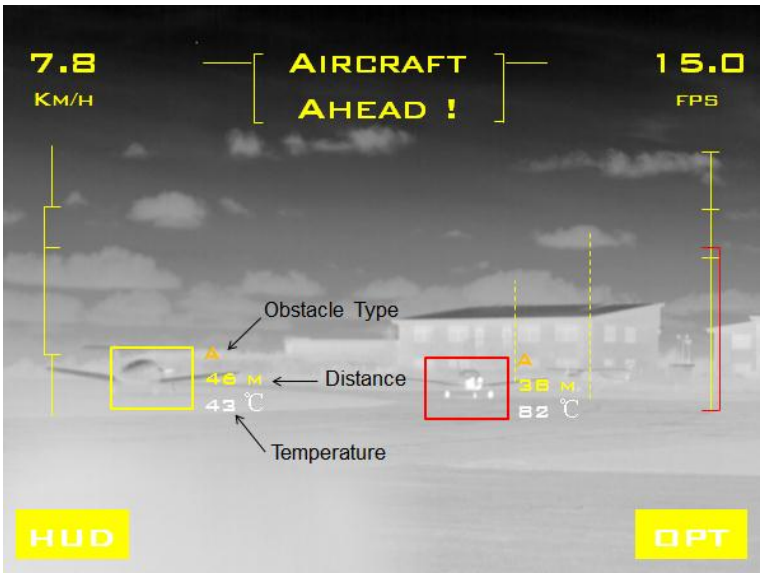


Figure 3-5: User Interface

Figure 3-6: HUD in Aircraft (image source[35])

3.5 Programming Language

All the algorithms implemented are programmed in Python. The code developed is included in the CD attached to this report.

The purpose of using Python is to make the code run faster, taking less time to process images. In [19], the system was coded in MATLAB which takes 10s to 30s to process each frame. To reduce the processing time, it is suggested that writing the code in a compiled language could be helpful. The system proposed in this thesis is supposed to run in real-time. Thus, in order to decrease the image processing time for each frame, Python is chosen as the programming language. Python has a highly optimized byte compiler and support libraries which make the code run fast. The proposed system is developed in Python 2.7.

Graphic library OpenCV 2.3 is used as the main library to process images. OpenCV is a computer vision library for real time application. It has C++, C, Python and Java interfaces running on different operation systems.

4 CAMERA MODEL AND CALIBRATION

4.1 Camera Model

A pinhole model is employed as the camera model in this thesis. A typical pinhole model is shown in Figure 4-1. In this model, the centre of the perspective projection is defined as the optical centre. The optical axis is the line perpendicular to the image plane and passes through the optical centre. The point in the image plane, which the optical axis goes through, is called the principal point. Among the parameters in the figure, P is the point of object, p is the object's point on imaging plane. f is the focal length and Z is the distance of object.

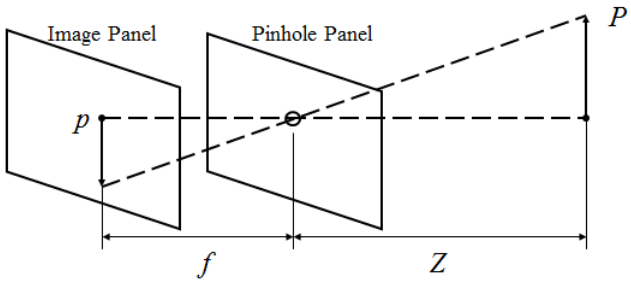


Figure 4-1: Pinhole Camera Model

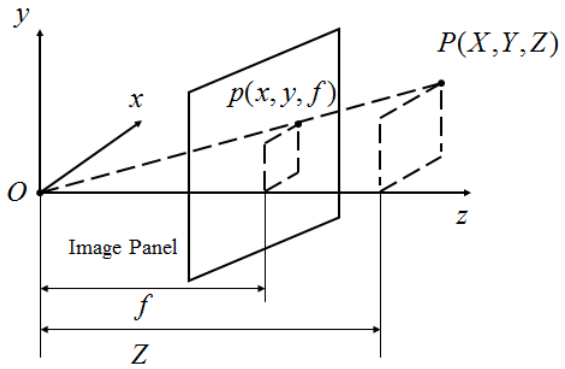


Figure 4-2: Pinhole Camera Model (Image panel shifted)

The principle of imaging employed in a pinhole camera is the perspective projection, which maps the object from 3D to 2D geometrically. It figured out the relationship between a 3D point and its 2D point on the image. The relation between the object point $P(X,Y,Z)$, focus length f, distance Z and image point $p(x,y,f)$ can be presented by similar triangles.

$$-\frac{x}{f}=\frac{X}{Z} \qquad \textbf{(4-1)}$$

By moving the image panel in front of the pinhole panel, the image is generated by the projection of the centre of the projection, as shown in Figure 4-2. Then, equation (4-1) can be written as

$$\frac{x}{f} = \frac{X}{Z} \tag{4-2}$$

The negative sign can be removed due to original image is upside down. Considering a camera with the camera model in Figure 4-2, $P(X,Y,Z)$ is the coordinates in real world and $p(x,y,z)$ is the pixel position on the image panel, the geometric relationship between P and p can be written as

$$x = \frac{Xf}{Z} \tag{4-3}$$

$$y = \frac{Yf}{Z} \tag{4-4}$$

The above equations assume that the length and width of pixel of the image sensor is equal. Then, the focus lens f for x and y are the same. However, this assumption is often invalid. In practice, the low-cost pixels of image sensors are often rectangular. Hence, focal length f_x and f_y are introduced for each direction. Let λ be the homogeneous scaling factor. Then the geometric relationship can be written as

$$(\lambda x, \lambda y, \lambda)^T = (Xf_x, Yf_y, Z)^T \tag{4-5}$$

Equation (4-5) can be the expressed in matrix notation,

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{4-6}$$

where $\mathbf{p} = [x, y, 1]^T$ is the point on the image panel and $\mathbf{P} = [X, Y, Z, 1]^T$ is the point on the object panel.

4.2 Intrinsic Parameters

4.2.1 Camera Coordinate System

The right-handed coordinate system is employed in the camera coordinate system. Figure 4-3 shows the right-handed coordinate system for camera. For an ideal camera, the optical centre is located at the same point as the origin of a 3D coordinate system. The axis X and axis Y are parallel to axis x and axis y of the image. The optical axis is collinear to the axis z and penetrates the image. The point in the image panel which optical axis penetrates is called principle point. For an image, the origin of the coordinate system is located on

the left-up corner of the image. Axis x and axis y are collinear to the width and height of the image. The coordinates are given in terms of pixel coordinates.

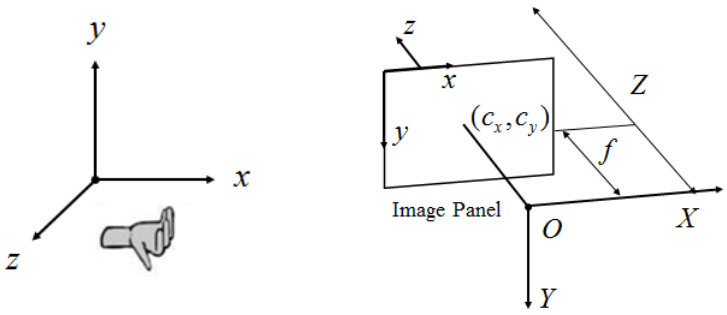


Figure 4-3: Camera Coordinate System

4.2.2 Principal-Point Offset and Image Skew

Usually, the centre of the projection (principle point) is not in the centre of the image of the camera, meaning the optical axis does not go through centre of the imaging sensor. Two parameters c_x and c_y are introduced to define the possible displacement of optical axis on the image.

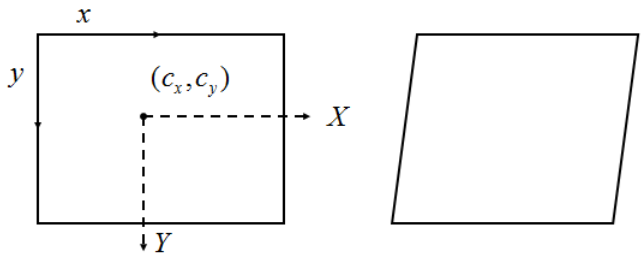


Figure 4-4: Principle Point offset and Image Skew

$$x = f_x\left(\frac{X}{Z}\right) + c_x \tag{4-7}$$

$$y = f_y\left(\frac{X}{Z}\right) + c_y \tag{4-8}$$

The pixel grid maybe skewed due to the synchronization problem of the image sampling process. Figure 4-4 (right) shows the skewed pixels. Parameter τ is used to define the skew of pixels.

$$x = f_x\left(\frac{X}{Z}\right) + \tau\left(\frac{Y}{Z}\right) + c_x \tag{4-9}$$

However, the skew of pixel grid is not considered in the calibration in this thesis, which means there is no problem in image sampling process.

Then, the perspective projection equation (4-6) can be updated as following

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \textbf{(4-10)}$$

where $\mathbf{p} = [x, y, 1]^T$ is the point on the image panel and $\mathbf{P} = [X, Y, Z, 1]^T$ is the point on the object panel.

4.2.3 Lens Distortion

Two kinds of distortions are considered in the camera model, the radial distortion and the tangential distortion. The radial distortion (shown in Figure 4-5) comes from the shape of lens. In general, the radial distortion can be rescaled according to the following equations.

$$x_{corrected} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \quad \textbf{(4-11)}$$

$$y_{corrected} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \quad \textbf{(4-12)}$$

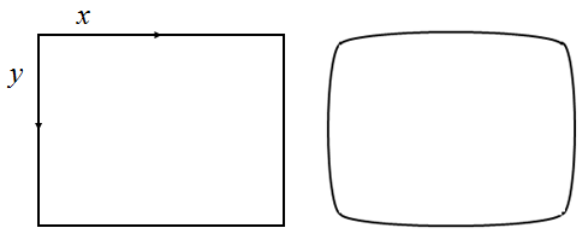


Figure 4-5: Lens Radial Distortion (image source: [36])

The tangential distortion (shown in Figure 4-6) is caused by the assembly process of the camera. The tangential distortion can be corrected by the following equations.

$$x_{corrected} = x + [2 p_1 y + p_2 (r^2 + 2x^2)] \quad \textbf{(4-13)}$$

$$y_{corrected} = y + [2 p_2 x + p_1 (r^2 + 2y^2)] \quad \textbf{(4-14)}$$

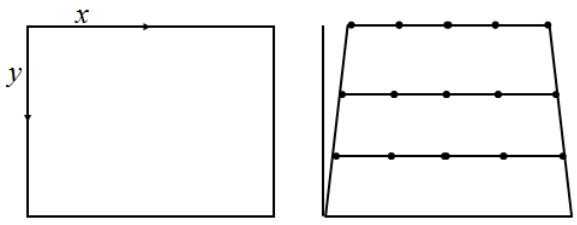


Figure 4-6: Lens Tangential Distortion (image source: [36])

4.3 Extrinsic Parameters

4.3.1 Rotation Matrix

Rotation Matrix indicates the rotation relationship between two coordinate systems, which is the coordinate system of camera and the 3D coordinate system of the world. For a 2D coordinate system, when the system rotates around axis z, the rotated point (x', y') can be written as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{4-15}$$

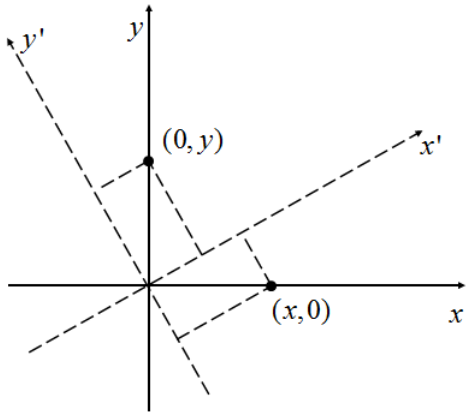


Figure 4-7: Rotate around One Axis (image source: [36])

For the 3D coordinate system, the rotation matrix $\mathbf{R}$ is the product of the three rotation matrices, $\mathbf{R}_x(\psi)$, $\mathbf{R}_y(\varphi)$ and $\mathbf{R}_z(\theta)$, which represent the rotation around each axis. ψ, φ and θ represent the rotation angles around axis x, y and z.

$$\mathbf{R}_x(\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\psi) & \sin(\psi) \\ 0 & -\sin(\psi) & \cos(\psi) \end{bmatrix} \tag{4-16}$$

$$\mathbf{R}_y(\varphi) = \begin{bmatrix} \cos(\varphi) & 0 & -\sin(\varphi) \\ 0 & 1 & 0 \\ \sin(\varphi) & 0 & \cos(\varphi) \end{bmatrix} \tag{4-17}$$

$$\mathbf{R}_z(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4-18}$$

Thus,

$$\mathbf{R} = \mathbf{R}_z(\theta)\mathbf{R}_y(\varphi)\mathbf{R}_x(\psi) \tag{4-19}$$

Then the rotation relationship of a point in two coordinate systems can be written as

$$\mathbf{P}_c = \mathbf{R}\mathbf{P}_o \tag{4-20}$$

4.3.2 Translation Vector

The translation vector is a shift from one coordinate system to another coordinate system. Figure 4-8 shows the translation of coordinate system. The relationship between two coordinate systems can be written as

$$\mathbf{P}_c = \mathbf{P}_o - \mathbf{T} \tag{4-21}$$

where $\mathbf{T} = [T_X \quad T_Y \quad T_Z]^T$.

The combination of rotation and translation of a coordinate system is shown in Figure 4-9. It can be writen as

$$\mathbf{P}_c = \mathbf{R}(\mathbf{P}_o - \mathbf{T}) \tag{4-22}$$

where $\mathbf{R}$ is the 3×3 rotation matrix and $\mathbf{T}$ is the 3×1 translation vector.

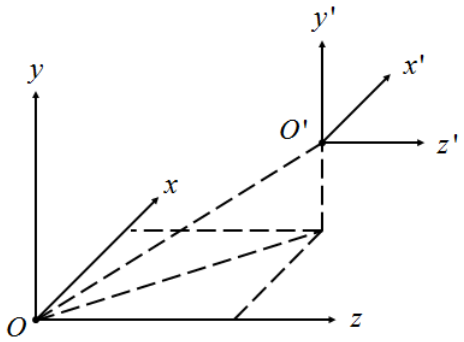


Figure 4-8: Translation of Coordinate System (image source: [36])

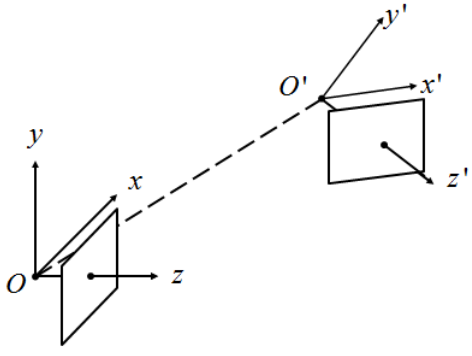


Figure 4-9: Rotation and Translation (image source: [36])

4.4 Camera Calibration

4.4.1 Homography

Homography relates the positions of the points on two planes, which is the object panel and the image panel in mono camera calibration. The transform of homography is shown in Figure 4-10. The action of the homography is defined as

$$\mathbf{p} = \lambda \mathbf{H} \mathbf{P} \tag{4-23}$$

where $\mathbf{p} = [x, y, 1]^T$ is the point on the image panel, $\mathbf{P} = [X, Y, Z, 1]^T$ is the point on the object panel and λ is a scale factor..

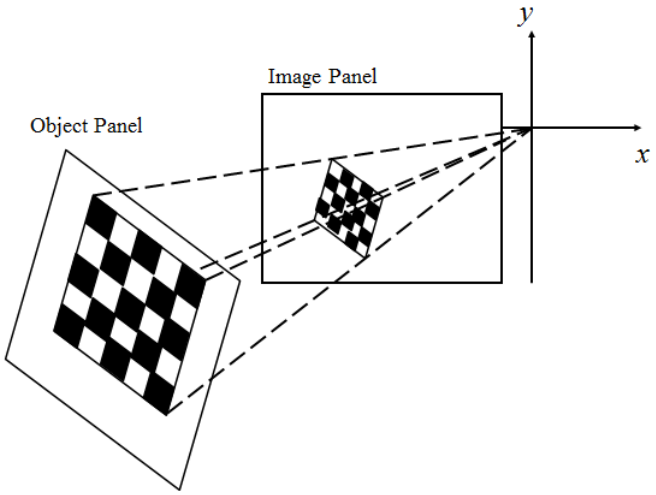


Figure 4-10: Homography (image source: [36])

The transformation can be expressed by another form which employ the rotation matrix $\mathbf{R}$, translation vector $\mathbf{T}$ and camera matrix. The effect of rotation matrix can be written as:

$$\mathbf{W} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix} \tag{4-24}$$

The camera matrix is

$$\mathbf{M} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{4-25}$$

Then, the relationship between $\mathbf{p}$ and $\mathbf{P}$ can be written as:

$$\mathbf{p} = \lambda \mathbf{M} \mathbf{W} \mathbf{P} \tag{4-26}$$

The equation (4-26) can be simplified if only one plan is concerned. By defining the object plane $Z = 0$, one of the columns of the rotation matrix become unnecessary. Then, equation (4-26) can be written as

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \lambda \mathbf{M} \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} \quad \textbf{(4-27)}$$

$$= \lambda \mathbf{M} \begin{bmatrix} r_{11} & r_{12} & T_x \\ r_{21} & r_{22} & T_y \\ r_{31} & r_{32} & T_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

As defined in equation (4-23),

$$\mathbf{p} = \lambda \mathbf{H} \mathbf{P} \quad \textbf{(4-28)}$$

where $\mathbf{H} = \mathbf{M} \begin{bmatrix} r_{11} & r_{12} & T_x \\ r_{21} & r_{22} & T_y \\ r_{31} & r_{32} & T_z \end{bmatrix}$ and $\mathbf{P} = \begin{bmatrix} X & Y & 1 \end{bmatrix}^T$.

Then, the positions of the points on two image planes are related by the homography matrix $\mathbf{H}$. Equation (4-28) is used in the camera calibration process. With multiple views of the chessboard, parameters can be estimated by calculating the homography matrix.

4.4.2 Camera Calibration

Single camera calibration is to calculate the intrinsic parameters of the camera. Much work has been done for optical camera calibration in computer vision. The main parameters which will be obtained in the calibration are the position of image centre in the image, focal length, different scaling factors for row pixels and column pixels, skew factor and lens distortion parameters. The algorithm which OpenCV used to solve for the calibration and lens distortion are based on Zhang [37] and Brown [38]. A chessboard is often used as the calibration pattern. The images of chessboard are captured from various viewpoints, shown in Figure 4-11. Points on the chessboard viewed through the camera can be regarded as a perspective transformation. The parameters for this transformation are contained in the homography matrix.

Figure 4-11: Images Captured in the Calibration

5 STEREO VISION

5.1 Fundamental Theory

5.1.1 Coordinate Systems

Figure 5-1 shows the stereo vision coordinate system. Two cameras are arranged frontal parallel and horizontally aligned. It can be seen the optical centre of the left camera is located at the origin of the coordinate system. The optical axis is collinear to axis z and goes through the left image. Axis x and axis y have the same direction as axis x_l and axis y_l in the image coordination system. The unit of the stereo coordinate system is pixel. The relation between the size of pixel and length is figured out in the stereo calibration. There exists a translation between the optical centres of both cameras. The translation can be described as a shift in axis x. In fact, the distance between two cameras is also called the baseline distance. The left and right images remain their own image coordination system.

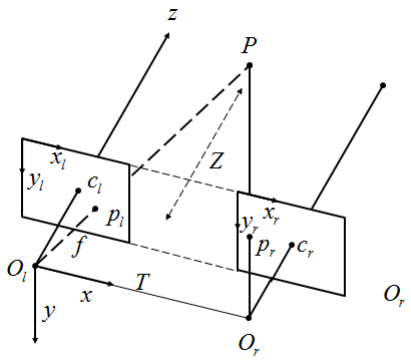


Figure 5-1: Stereo Vision Coordinate System (image source: [36])

5.1.2 Triangulation Principle

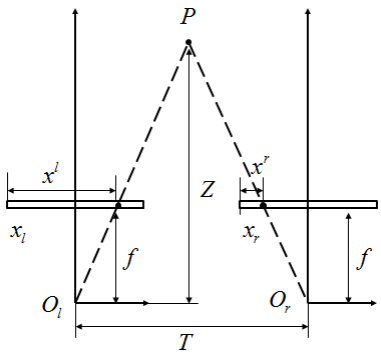


Figure 5-2: Triangulation Principle (image source: [36])

A basic principle in the stereo vision is the triangulation principle. Obstacle distance is estimated by the triangulation principle. The distance measuring using triangulation principle is shown in Figure 5-2. It can be seen two cameras

capture the same scene from two different viewpoints. Assuming the two cameras are aligned and not distorted, then the image planes of two cameras are exactly coplanar with each other. Two points (O_l and O_r) are the centre of projection while T is the distance between the two points. The optical axes (the ray from the centre of projection) are exactly parallel. The focal lengths (f_l and f_r) of the two cameras are equal. Then, in the left and right image, principal points C_x^{left} and C_x^{right} have the same pixel coordinates. x^l and x^r are the horizontal positions of the points in the left and right image. The disparity of point P in the right and left image can be calculated by $d = x^l - x^r$. Using similar triangles, the depth Z is given as following

$$\frac{T-(x^l-x^r)}{Z-f}=\frac{T}{Z}\Rightarrow Z=\frac{fT}{x^l-x^r} \qquad \textbf{(5-1)}$$

Thus, once f , T and disparity d are known, the depth Z can be obtained.

5.1.3 Epipolar Geometry

Epipolar geometry is used to constrain the 3D point in view of the cameras when the two cameras are not aligned. In real situation, the two cameras are not frontal parallel and horizontally aligned. Then, the imagers captured are not parallel to each other. However, when two cameras are viewing the same scene, the two images are constrained by the epipolar geometry. Figure 5-3 shows the principle of epipolar geometry. Objects in line O_lP linking left camera's focal point O_l and P can be seen in the right image. The correspondence for a point belonging to line O_lP lies on this line. This line is called the epipolar line. When matching point P in two images, only the epipolar line corresponding to this point in the search image need to be searched. It turns the 2D searching problem to a 1D line searching problem, which is of great importance in the correspondence. Much computation can be saved when the searching space is limited to 1D. Moreover, the order of points is also constrained. The same order of points will occur in both the left and the right images. This can be used to judge whether or not the correspondence is correct.

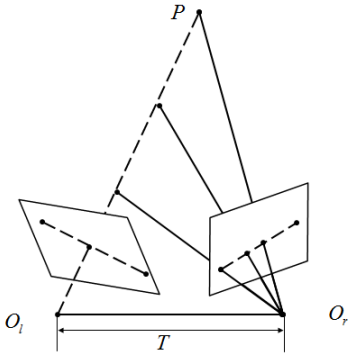


Figure 5-3: Epipolar Geometry (image source: [36])

5.1.4 Essential and Fundamental Matrices

The essential matrix describes the translation and rotation information between the left and the right camera. It relates the location of two cameras in physical coordinates which is geometrical. The essential matrix can be written as

$$\mathbf{E} = \mathbf{RS} \tag{5-2}$$

where

$$\mathbf{S} = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & T_x \\ -T_y & T_x & 0 \end{bmatrix}.$$

The relation between the points in left image and right image is

$$\mathbf{p}_r^T \mathbf{E} \mathbf{p}_l = 0 \tag{5-3}$$

Fundamental matrix indicates the image relationship between the left and the right camera in pixel coordinates. The information of essential matrix is included in the fundamental matrix. Additionally, it also has the information of the intrinsic parameters of both cameras. The fundamental matrix can be written as

$$\mathbf{F} = (\mathbf{M}_r^{-1})^T \mathbf{E} \mathbf{M}_l^{-1} \tag{5-4}$$

where

$$\mathbf{M} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}.$$

Also, the relation between the points in left image and right image is

$$\mathbf{p}_r^T \mathbf{F} \mathbf{p}_l = 0 \tag{5-5}$$

5.2 Stereo Calibration

When using stereo vision, the spatial relation between two cameras needs to be known and distortion due to the imperfections of the lens needs to be corrected. The purpose of stereo calibration is to calculate the geometrical relationship between two cameras in space. The parameters and matrices obtained during the calibration are lens distortion parameters, intrinsic parameters, rotation matrix $\mathbf{R}$, translation vector $\mathbf{T}$, the essential and fundamental matrices. These parameters will be used to rectify the image in the process of stereo rectification and to calculate the disparity map in the process of stereo correspondence.

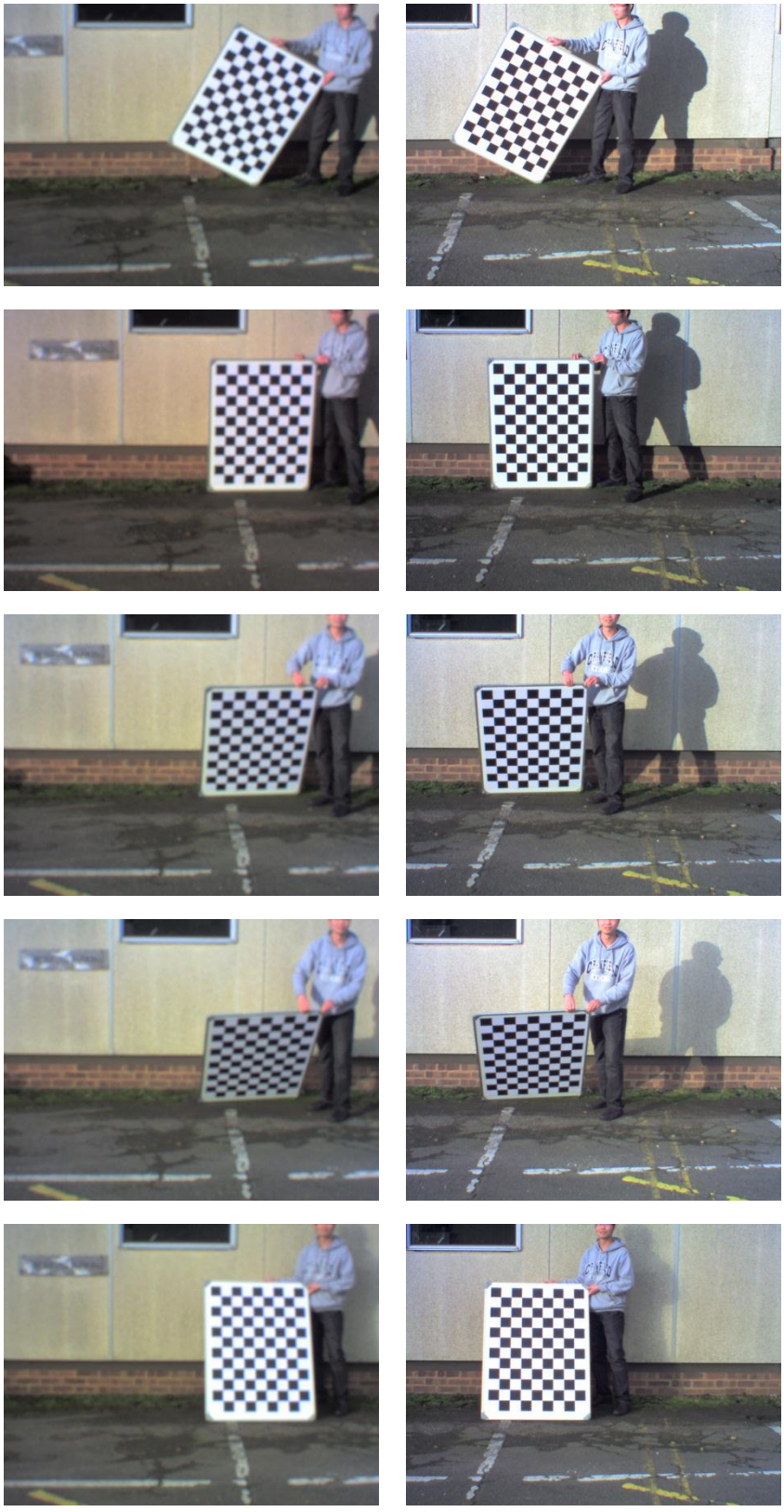


Figure 5-4: Images Captured by Two Cameras

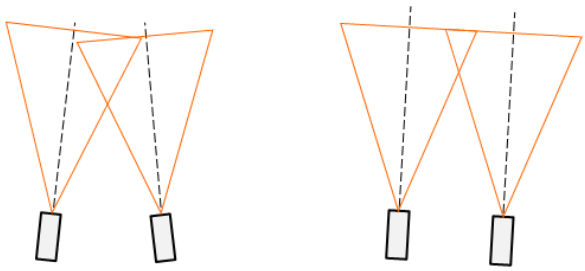
Figure 5-4 shows the images captured by two optical cameras. The same scene was captured by the two cameras. The left colum shows images captured by the left camera while the right colum shows images captured by the right camera.

When calibrating the stereo system, it is better to place the two cameras frontal parallel and horizontally aligned physically. Although the rectification will align the two cameras mathematically, this will make the mathematical transformations more tractable. Otherwise, extreme image distortions can be created by the mathematical alignment. Also, the stereo overlap area of the captured images will be reduced or eliminated.

The parameters obtained in the single camera calibration can be used in the stereo calibration for an initial value.

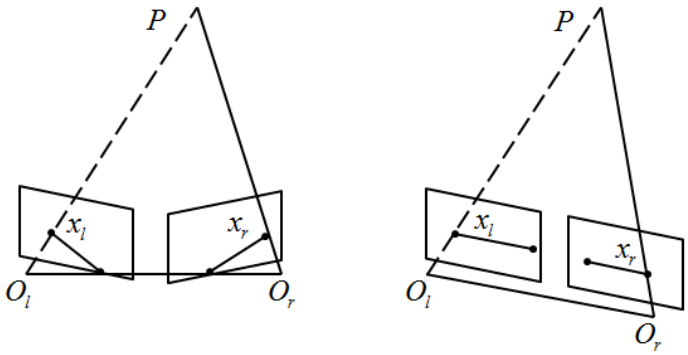
5.3 Stereo Rectification

In reality, the optical axes of the two cameras are not parallel to each other (Figure 5-5 (a)). Therefore, images captured by both cameras are not on the same plane (Figure 5-6 (a)). The purpose of rectification is to align the two cameras into one viewing plane mathematically (Figure 5-5 (b)). It will transform the images on which the epipolar lines are parallel and horizontal (Figure 5-6 (b)). The whole process is based on the epipolar geometry.



(a) Cameras in real world **(b)** Cameras aligned mathematically

Figure 5-5: Align the Two Cameras into One Viewing Plane Mathematically



(a) Images before rectification **(b)** Images after rectification

Figure 5-6: Transform the Left and Right Images in One Plane (image source: [36])

Figure 5-7 shows the principle of rectification. The left and right image planes are not on the same plane. Given a plane in space, there exists two homographies that map each image onto plane H. If plane H is parallel to line O_lO_r, then epipolar lines in the pair of rectified images are parallel. With proper choice of the coordinate system, epipolar lines can be parallel to the row of the image. The algorithm for rectification is as follow:

1. Select a plane H parallel to O_lO_r
2. Define the left and right image coordinate systems on H
3. Construct the rectification matrices and the rectified image's coordinate systems.

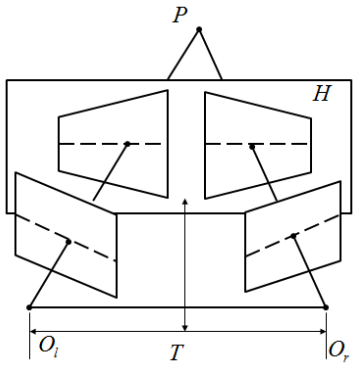


Figure 5-7: Principle of Rectification

5.4 Stereo Correspondence and Re-projection

Stereo correspondence is to find finding homologous points in the stereo pair. Figure 5-8 show the processing of matching points x^l and x^r in the left and the right image. This processing is carried out after image rectification. Once the pixels and the corresponding pixel location are matched, the disparity map can be calculated. Then, a disparity map can be obtained by calculating the whole matched pixels. In [39], the algorithm flows for stereo correspondence is described.

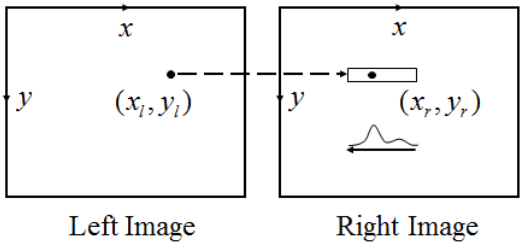


Figure 5-8: Correspondence (image source: [36])

The correspondence is a challenging work because a number of factors will affect the result, such as specular surface of objects, uniform regions, and transparent objects. The algorithms used in OpenCV to correspondence are Graph Cut, Block Matching and Semi Global Block Matching. Among these algorithms, blocking matching has the least computation cost. However, the

correspondence quality of block matching is inferior. Since the proposed systems have real-time performance requirements, the stereo correspondence routines need to run fast. Thus, the block matching algorithm is adopted as the correspondence method.

Once the disparity map has been obtained in correspondence, the three dimensions coordinates of points in world can be calculated by re-projecting the coordinates of points on the disparity map. The 3D depth map can be obtained by equation (5-6), where $p(x, y)$ is the given point and d is the disparity $d = x_l - x_r$.

$$\mathbf{Q}\begin{bmatrix} x \\ y \\ d \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix} \tag{5-6}$$

where

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -c_y \\ 0 & 0 & 0 & f \\ 0 & 0 & -1/T_x & (c_x - c_x')/T_x \end{bmatrix} \tag{5-7}$$

In equation (5-7), the calculation is based on the left camera except for c_x' . c_x' is the principal point in the right image and $c_x = c_x$' under the condition of the principal rays of two cameras intersect at infinity. Then, the 3D coordinates of P in real world is

$$P = (\frac{X}{W}, \frac{Y}{W}, \frac{Z}{W}) \tag{5-8}$$

6 OBSTACLE DETECTION

6.1 Temperature Measurement

Unlike the other temperature measurement methods such as thermocouple, infrared camera measures the temperature by capture infrared radiation. It receives radiation from the object itself, the surrounding radiation reflected via the object surface and the atmosphere. Figure 6-1 shows the general thermal graphic measurement situation. It can be seen part of the infrared radiation is absorbed by the atmosphere during transmission. In outdoor measurements, both reflection of sunlight and absorption by atmosphere make the temperature measured shift from its actual temperature. The received radiation power is

$$W_{total} = \varepsilon\tau W_{obj} + (1-\varepsilon)\tau W_{reft} + (1-\tau)W_{atm} \quad \textbf{(6-1)}$$

where W is the radiation power, ε is the emittance of object and τ is the transmittance of the atmosphere.

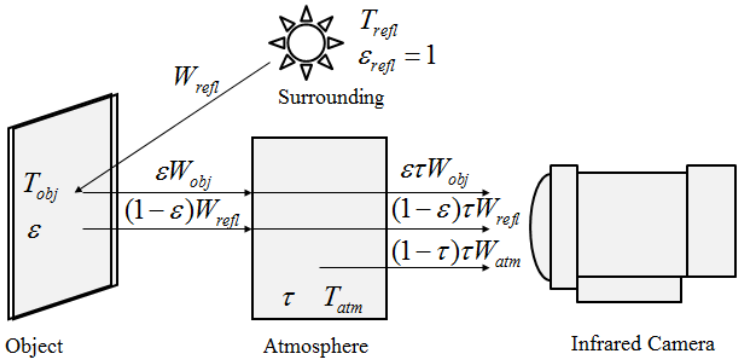


Figure 6-1: Temperature Measurement of Infrared Camera (image source: [40])

In real conditions, there are often many objects on the scene. Emittance for all the objects cannot be defined individually. Also, for a moving object, the temperature measured also varies according to the distance because of the absorption of radiation power by atmosphere. However, in this application, the accuracy of temperature measurement is not of great importance. Image quality (such as contrast between object and background) is more important than the actual temperature of objects.

6.2 Image Processing

6.2.1 Image Normalization

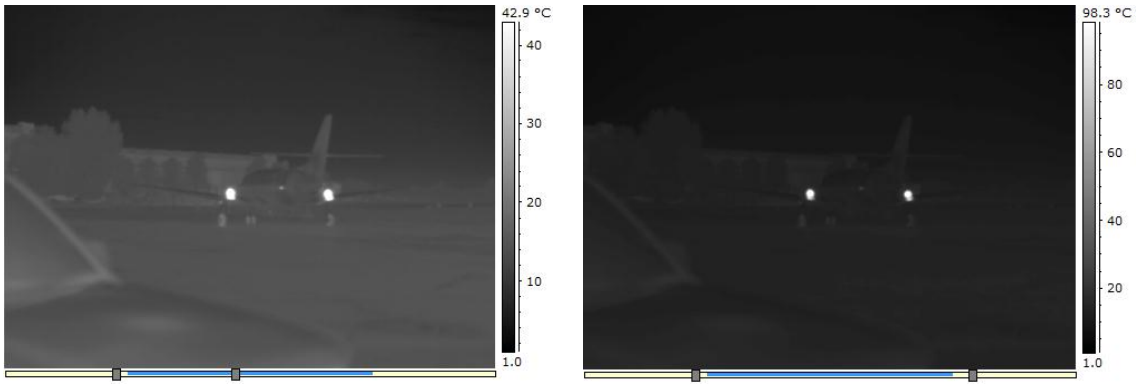
Image normalization is to transform the captured temperature matrix into a grey level image. Originally, the thermal data file captured by IR camera contains all the temperature information. These data are visualized by transforming them into 8-bit grey level image, which has a value range from 0-255. Temperature information may be lost in the transformation. Figure 6-2 shows the normalized infrared image with different temperature scales which are 0~42.9 ℃ and 0~98.3℃. It can be seen from the images, the minimum and maximum value of

temperature will influence the visibility of image. Temperature information above 42.9 ℃ will be lost If the temperature scale is set as 0~42.9 ℃ in the normalization (shown in Figure (a)). The normalization is written as

$$G = \frac{T_m - T_{low}}{T_{high} - T_{low}} \times 255 \quad \textbf{(6-2)}$$

where T_m is the measured temperature, G is the grey level, T_{high} and T_{low} are margin of the temperature scale.

Since different type cameras have their own methods to normalize the image, for instance, the FLIR T400 takes the lowest and highest temperature as the temperature scale while the FLIR SC640 take the lowest and the average temperature as the temperature scale. Thus, sometimes it is necessary to normalize the image such that the temperature scale contains the lowest and highest temperature.



(a) Thermal image normalized with temperature 1~42.9℃

(b) Thermal image normalized with temperature 1~98.3℃

Figure 6-2: Infared Image with Different Temperature Scale

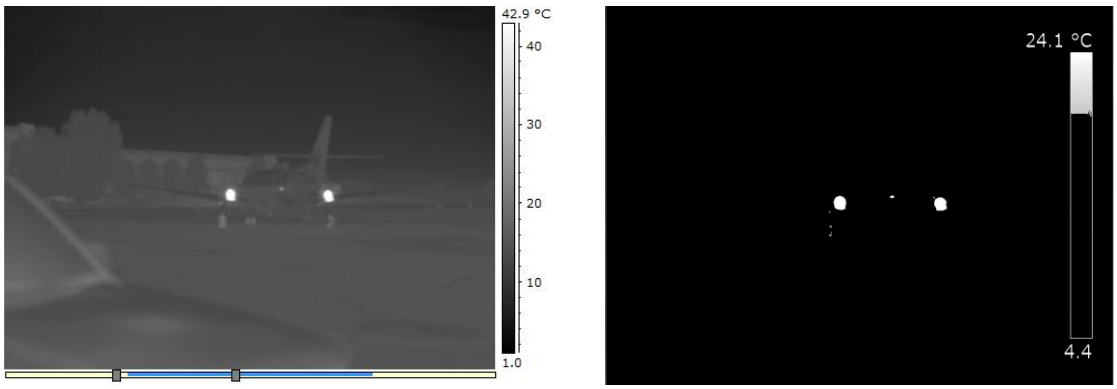
6.2.2 Background Elimination

Background is an issue in optical stereo vision systems because it may be regarded as obstacle. In the optical stereo vision system developed in [41], the ground is eliminated by assuming the ground is flat and a height threshold is set to ignore the obstacle above the height threshold.

Since the background has a relatively low temperature in thermal images, the background can be eliminated by setting the threshold of temperature. Once the temperature range of obstacle is known, a threshold of the temperature can be set up to eliminate the background. The threshold of temperature needs to be transformed to a grey value based on equation (6-2) providing the temperature scale of the grey image is given.

Figure 6-3 shows the original thermal image (Figure (a)) and background elimination thermal image (Figure (b)). It can be seen in the original thermal image, there are buildings, trees, airplane and ground in the same scene. By setting the threshold of temperature (temperature will be converted into grey

value in the processing), the ground, trees and building are eliminated. Only two engines are left in the right image.



(a) Original thermal image **(b)** Background eliminated image

Figure 6-3: Background Elimination

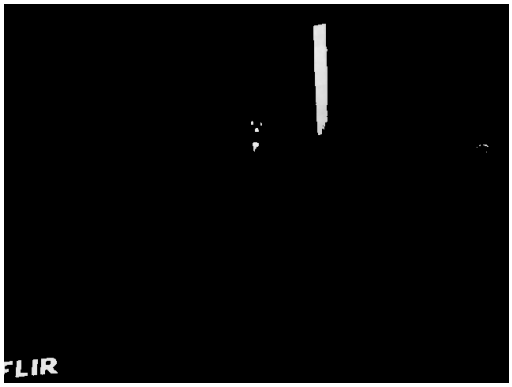
The range data of the low temperature obstacle may be lost in the background elimination. In Figure 6-3, since only the engine has a high temperature, the outline of other parts of the airplane is also removed from the image. The consequence is the distance of other parts such as tail is lost. If the distance of engine is taken as the distance of the whole airplane, the actual distance of the whole airplane is shorter than the distance measured. However, the original thermal image will be provided to the pilot in order to reduce the negative influence of back ground elimination.

Figure 6-4 shows the depth maps calculated from the original images (left column) and back ground eliminated images (right column). Figure (a) is the original rectified thermal image from left camera of the stereo vision system while Figure (b) is the background eliminated and rectified image. Figure (c) and Figure (e) are the depth maps calculated from original rectified image. Figure (d) and Figure (f) are the depth maps calculated from background eliminated thermal image. The depth maps in Figure (c) and Figure (d) are viewed in XY plane. The depth maps in Figure (e) and Figure (f) are viewed in XZ plane.

Figure (c) and (d) show there are less contours in the depth map computed from background eliminated images. This means less contours will be involved in the depth map filtering.

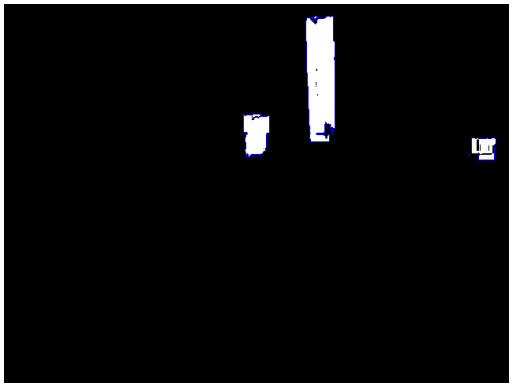
However, it is impossible to eliminate all the background by setting the temperature threshold of the image. In Figure (a), there are building, car, trees and people in the thermal image. In Figure (b), people and building are left in the image. The building remains after background elimination because the temperature of building is relatively high due to the reflection of sunlight. Then, the building is regarded as an obstacle and distance of building is estimated in Figure (d). Obviously, the range data of building needs to be filtered.

(a) Original Rectified thermal image

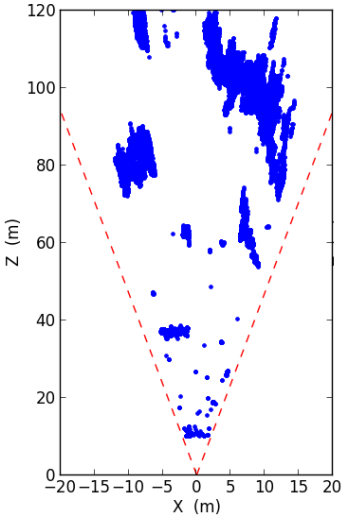


(b) Background eliminated

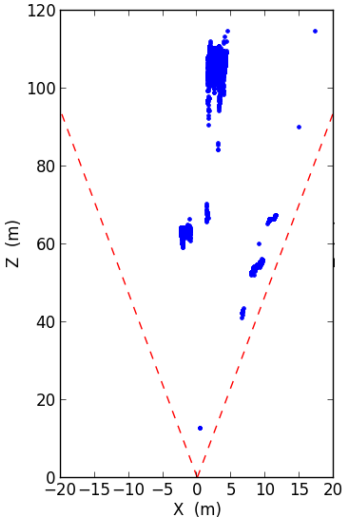
(c) Depth map calculated from Figure (a)



(d) Depth map calculated from Figure (b)



(e) Depth map calculated from Figure (a) (view in XZ plane)



(f) Depth map calculated from Figure (b) (view in XZ plane)

Figure 6-4: The Influence of Background Elimination

6.2.3 Depth Map Filtering

There are noises in the depth map due to mismatching or unwanted obstacles which have been discussed in Chapter 6.2.2. In Figure 6-5 (b), it can be seen in the original depth map, there are some points which do not belong to people or building. They are the mismatching points in the corresponding process. Since the contour area in the depth map can be calculated, noise (contour with very small area) can be filtered by setting a threshold of contour area. As discussed in Chapter 3.2, assuming the minimum area can be identified by pilots is 10X10 pixels, a contour area of 100 pixels can be used as a threshold to filter these small area contours.

Figure 6-5 (a) shows two contours which represent people and buildings. Obviously, the building is not the obstacle of interest. The filtering of large contour area such as buildings is based on a combination of contour area and distance. Once the contour area and distance are calculated, a rough size of obstacle can be calculated. Assuming the shape of obstacle is a square, the length of a side can be estimated by equation (6-3),

$$l_s = \frac{f}{D} L \tag{6-3}$$

where l_s is the length of a side of the square, L is the length of object, D is the distance, f is the focus length. Then the rough size of an object is

$$s = \left(\frac{f}{D} \right)^2 S \tag{6-4}$$

where D is the distance, s is the area of contour in image, S is the size of obstacle. By setting the object size S , contours which have an area bigger than s can be removed in the depth map.

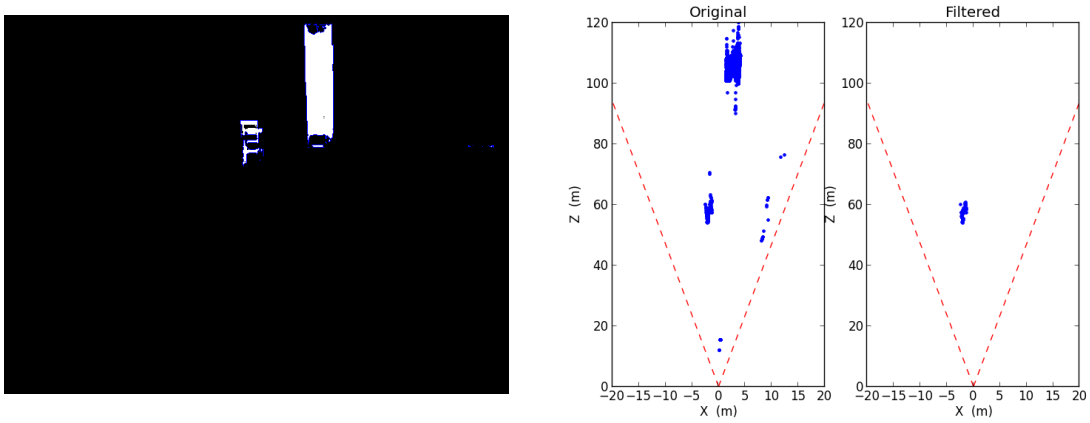
For instance, the engine size of A380 is 2.94 m (Rolls-Royce Trent 900, fan tip) and 2.95 m (Engine Alliance GP7000, fan tip) in diameter. Then, the section area of engine is about 6.83 m^2. Assuming a camera has an angle of view $25^o \times 19^o$, detector size $8mm \times 6mm$ and sensor resolution $640 \times 480 pixels$, then the focal length is $1443 pixels$, the relation between pixel and length is $80 pixels / mm$. At the distance of 40 metres, s is $1.6232\text{e-}006 m^2$ (about $10388 pixels$). At the distance of 80 metres, s is $4.0580\text{e-}007 m^2$ (about $2597 pixels$). At different distance, object with contour area larger than s can be removed.

The algorithm flow of depth map filtering is as follow:

1. Find contours in the depth map
2. Calculate contour area and get distance information for each contour
3. Calculate s according to the distance and maximum area of object predefined

4. Select the contours and extract the coordinates of contours

Figure 6-5 (b) show the depth map filtered based on contour area proposed. The left image shows the original depth map. It can be seen there are range points of people, building and noise in the depth map. The right image shows the filtered depth map. It can be seen the points of building and noise have been eliminated.



(a) Depth map (viewed in XY plane) **(b)** Depth map filtering (viewed in XZ plane)

Figure 6-5: Depth map filtered based on contour area

6.3 Obstacle Classification

6.3.1 Obstacle Temperature

Different kinds of objects can be found at the airport, including all kinds of airplanes, helicopters, trucks, cars and people. In this thesis, objects are divided into three main categories, people, vehicle and airplanes. A simple way to classify the obstacle is by setting a temperature range for each obstacle. For a given temperature, obstacles can be classified into different categories based on the temperature range. This method works well when there is a temperature gap between obstacles, such as airplanes and vehicles. The airplane can be easily identified because of high temperature. However, when there is overlap in temperature range, it is difficult to classify the obstacle. For instance, given temperature 31℃, the obstacle could be a car or a people.

A test was carried out in order to measure the temperature range of different objects. Table 6-1 shows the temperature range of obstacles measured at Cranfield Airport. Temperature is measured by FLIR T400 and FLIR SC640. Since the highest detecting range of the FLIR T400 is set as 150℃, the actual temperature may be higher than the numerical reading. It can be seen most of the airplane measured are small airplanes. For large airplanes such as jets, the temperature of the engine would be much higher than smaller airplanes measured in the table. Also, the tests were carried out under different weather conditions.

Table 6-1: Temperature range of obstacles

	Measured Temperature Range (℃)					
Weather	**People**	**Car**	**Truck**	**Small Twin Propeller**	**Bulldog**	**Helicopter**
Sunny	22~31.8	30~49.5	30.5~123.5	107.3~134.7	51.4~165.8	88~216.4
Rainy	20~33.3	34	33.8~71.6	69~168	39.8~133.4	139~435
Windy	24.9	22.5~33.9	23.8~25.9	86.9~150		190.4
Foggy	16.2~28	51~54				
Cloudy		41~44				

6.3.2 Fuzzy Logic

Fuzzy logic is a form of probabilistic logic. The variables in a fuzzy logic system have a truth value that ranges in degree between 0 and 1 while there are only two fixed values (0 and 1) in a traditional binary logic system. Fuzzy logic has been widely applied in fields such as process control, pattern recognition and object classification. It also has been used in image classification. In [42], fuzzy logic is employed to classify the land cover. The land cover classes are classified based on the spectral information of objects. Three bands of the spectral which are green, red and NIR are used as inputs of the fuzzy logic system. In [43], RGB components of pixels are used as inputs of the fuzzy logic system in satellite image classification. Not only RGB colour space but also other colour space can be used in the fuzzy logic system. In [44], the YCbCr colour space is employed in a colour model for fire pixel classification. It aims to separate the luminance component for fire detection. Thus, for the application of thermal image classification, the temperature can be used in fuzzy logic.

As discussed in Chapter 6.3.1, objects can be classified based on the temperature range of objects. Apparently, it is difficult to set temperature ranges for objects which have temperature overlaps. However, the employment of fuzzy logic can make the classifier more general and robust by turning the classification into a probability problem instead of simply setting a temperature range. The temperature will be converted into a truth value. The classification is accomplished by comparing the values calculated by different membership functions.

For the real application, the measured temperature distributions of objects need to be calculated first, shown in Figure 6-7. Three types of membership functions are employed, which are Gauss (for people), Sigmoid (for helicopter) and Gbell (for the rest). The definition of the three membership functions are shown in equation (6-5), (6-6) and (6-7). Second, membership functions are built according to the temperature distribution of obstacles. The same measured

temperature can be mapped by the function a truth value in the 0 to 1 range. These truth values represent the probability of obstacle type. The setting of each membership function is shown in Table 6-2. Figure 6-6 shows the plots of the membership functions.

Gauss:

$$u(x) = e^{-\frac{(x-c)2}{2\sigma^2}} \tag{6-5}$$

Gbell:

$$u(x) = \frac{1}{1+\left|\frac{x-c}{a}\right|^{2b}} \tag{6-6}$$

Sigmoid:

$$u(x) = \frac{1}{1+e^{-a(x-c)}} \tag{6-7}$$

Table 6-2: Membership function Settings

People	Gauss	$\sigma = 3, c = 27$
Car	Gbell	$a = 10, b = 4, c = 40$
Truck	Gbell	$a = 15, b = 4, c = 60$
Bulldog	Gbell	$a = 17, b = 4, c = 100$
Twin-propeller	Gbell	$a = 20, b = 4, c = 130$
Helicopter	Sigmoid	$a = 0.2, c = 140$

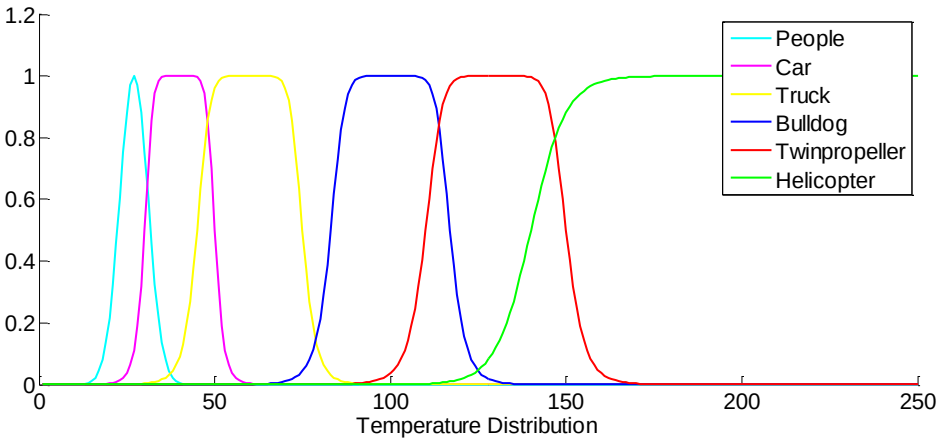


Figure 6-6: Membership Functions

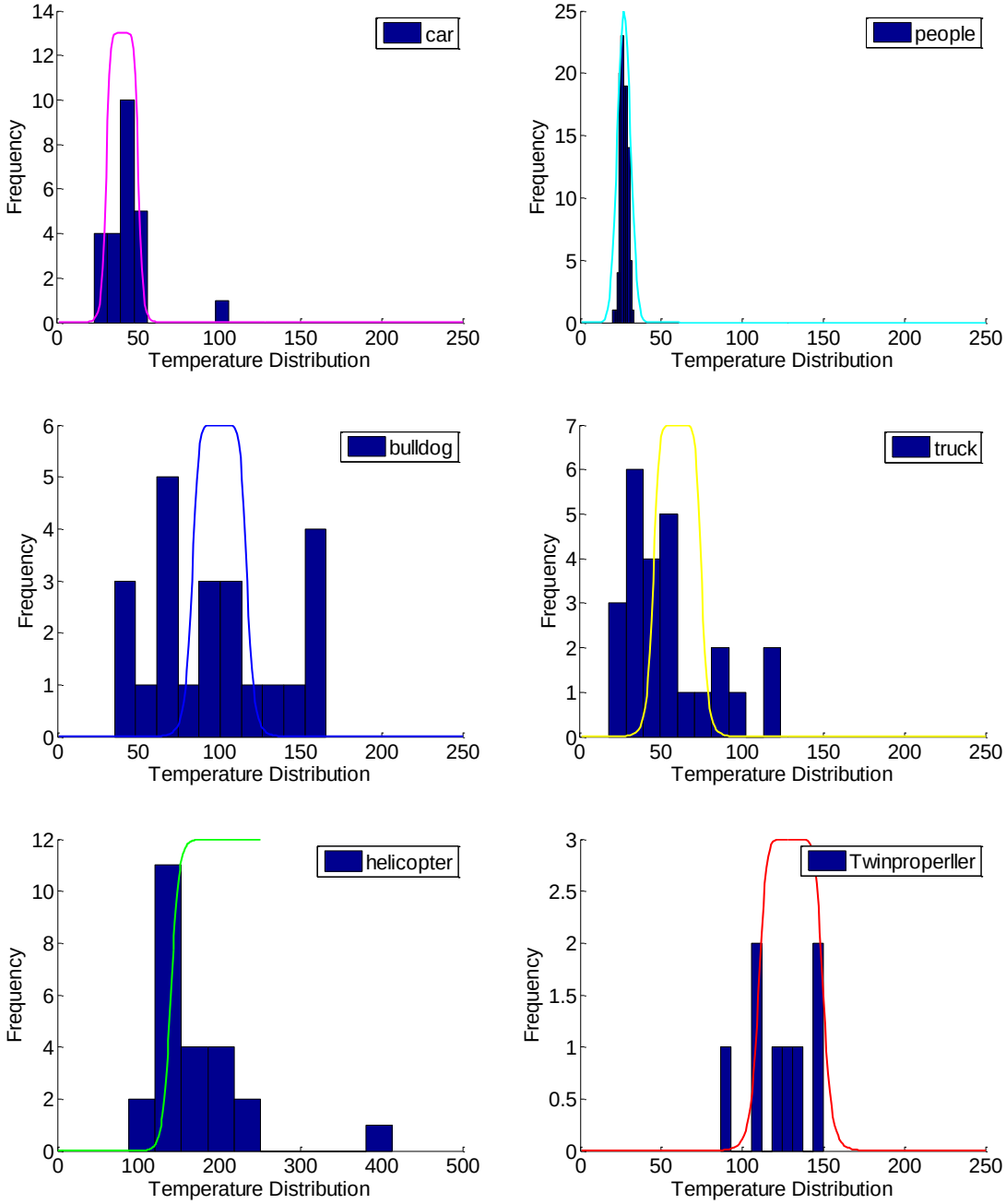


Figure 6-7: Measured Temperature Distribution of Obstacles

6.3.3 Position Involved

It can be seen in Figure 6-6 that there exit overlaps in temperature range which make the system hard to classify the type of obstacle. A concept of using the position information is proposed to assist obstacle classification. Assuming the proposed system is mounted on the aircraft, the position information of the aircraft (such as GPS information) can be used to identify the area of airport. Possible obstacles can be predefined on that area. For instance, when the aircraft is preparing taking-off on the run-way, the probability of obstacle to be a people or a truck is slim. Then, even if the temperature detected is low, the

obstacle is more likely to be an aircraft. In the algorithm, when comparing the truth values, there is no need to take people and trucks into consideration. Different areas of the airport can be divided into three categories, run-way, taxi-way and apron. Only aircraft are considered on the run-way while both vehicle and aircraft are considered on the taxi-way. For apron, all obstacle types need to be taken into consideration.

The obstacle classification based on the temperature aligned with position information is just a concept. In the future work, other algorithm can be used to classify the obstacle. For instance, the pattern recognition technique in [45] or the contour model of human in [9] can be employed to identify people. However, the type of obstacle provided by the system is just a reference for the pilot. What pilots see in the thermal image can give them more information. The type of obstacle can be further confirmed by the pilots' own judgement.

6.4 Algorithm Flow

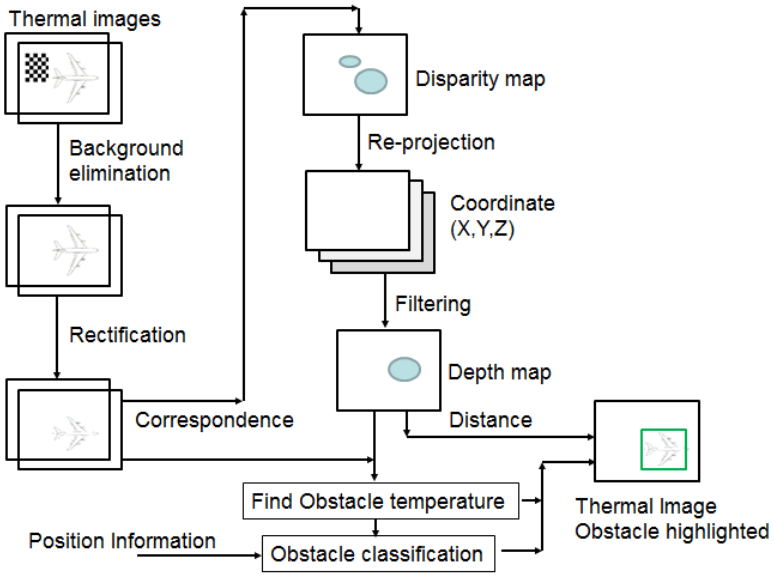


Figure 6-8: Algorithm Flow

Figure 6-8 shows the algorithm flow of the proposed system. First, images are captured by both left and right camera simultaneously. The backgrounds of the left and right image captured at the same time are eliminated according to the temperature threshold. Next, both images are rectified using parameters calculated in the stereo calibration. After that, correspondence is used to matching the points in the both rectified images. A disparity map is obtained by the process of correspondence. Next, a 3D depth map can be obtained by re-projection of the disparity map. The 3D depth map is then split into three planes, which are X plane, Y plane and Z plane. The Z plane is used in order to get the distance of the obstacle and find the contours of objects. After filtering the depth map (Z plane) based on contour area and distances, obstacles of interest are left. Next, the obstacles' temperature in the left rectified thermal image is extracted to carry out classification. The obstacle type can be classified according to the obstacle temperature and system position information. Finally,

the bounding of obstacle is highlighted by a rectangle with obstacle type, distance and temperature labelled beside. A warning will pop up if the distance is less than 40 metres.

The image processing speed is evaluated by processing 50 frames of thermal images. The result is shown in Figure 6-9. It can be seen the minimum processing time is about 0.19s, the maximum processing time is about 0.54s and the average processing time is about 0.43s.

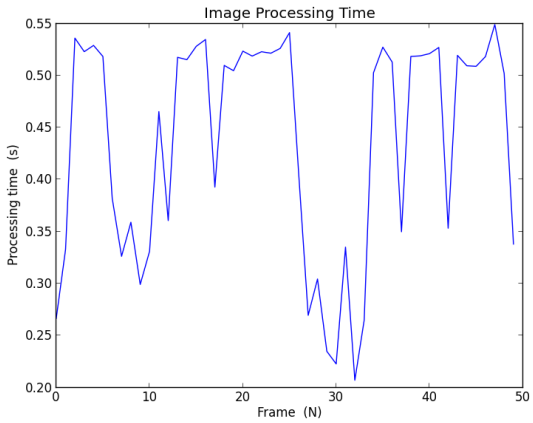


Figure 6-9: Image Processing Time

Figure 6-10 shows the image processing time calculated from the same images without background elimination. It can be seen the average processing time in Figure 6-10 is less than the average processing time in Figure 6-9 . After background elimination, the low texture of image will cause the matching algorithm more time to calculate.

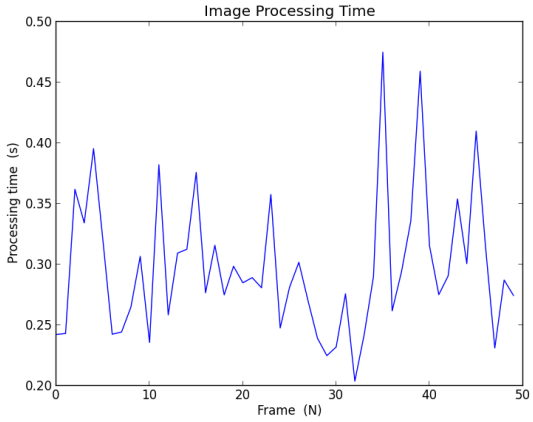


Figure 6-10: Image Processing Time (without background elimination)

7 SYSTEM TESTING

7.1 System Setup

7.1.1 Hardware Setup

(a) FLIR ThermCAM SC640

(b) Stereo vision system with two thermal cameras

Figure 7-1: Stereo Vision System with a Pair of Infrared Camera

Figure 7-1 shows the hardware of the whole system. The type of infrared camera is FLIR ThermCAM SC640. Two infrared cameras are employed in the system testing. Cameras are mounted on a steel frame supported by a tripod. The steel frame has a flat surface which keeps the two cameras on the same plane for the purpose of keeping their optical rays parallel to each other. It is necessary to make the optical rays of the two cameras as paralleled as possible since less distortion of images will be caused in the rectification step if two cameras are physically parallel to each other.

There are three types of lens for selecting, which are 45°x 49°, 24° x18° and a micro-lens. In order to have a long detecting range, a lens with angle of view 24°x18° is chosen. The frame rate of this type of camera can reach 30Hz. However, in the real test, 15Hz was selected because one of the cameras didn't work well at the frame rate of 30Hz. The detector type of SC640 is Focal Plane Array (FPA), uncooled micro-bolometer 640x480 pixels. There are three temperature ranges can be selected which are -40~120℃, 0~500℃ and 300~2000℃. In the real test, the temperature range was selected according to the real situation. The two cameras had the same setting before testing. The main parameters of the camera are set as follows:

1. Emission Rate: 0.95
2. Relative Humidity: 50%
3. Atmosphere temperature: 20 ˚C
4. Temperature Range: -40~120 ˚C or 0~500℃ (depend on the situation)
5. Distance between thermal infrared camera and Obstacle: Estimated

Since material such as steel, aluminium and the skin of people have a different emission rate, the value chosen here does not represent any material. The distance between two cameras is set about 0.84m. Images captured by thermal

infrared cameras were processed in a laptop. The specification of the laptop is shown in Table 7-1.

Table 7-1: Computer Specification

Processor	Intel Core Duo L9400@1.86Ghz
RAM	4 GB DDR3 1067MHz
Hard Drive	Kingston SSD 128G
GPU	Intel X4500 Integrated Graphics
Operating System	Windows 7 (32 bit)

7.1.2 Baseline Distance

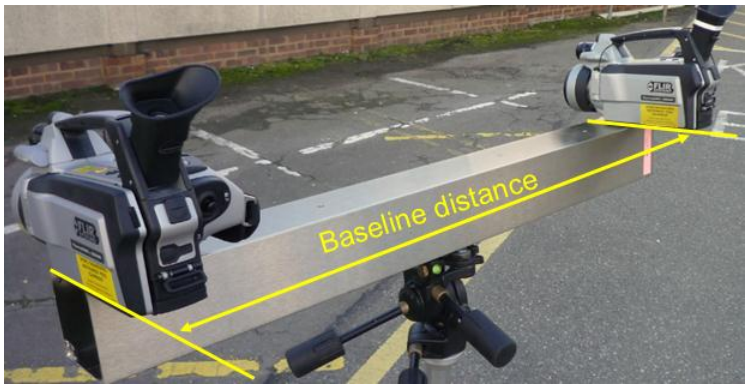


Figure 7-2: Baseline Distance

The baseline distance is the distance between two cameras in a stereo vision system. The choosing of baseline distance is based on the detecting range and range resolution. In equation (4-2), the depth is inversely proportional to disparity. A nonlinear relationship exists between depth and disparity. When objects are far away from the camera, large depth differences make for small disparity differences. When objects are near the camera, small depth differences will result in a great disparity. This means the closer the objects are to the camera, the stereo vision system will have a higher depth resolution. For a required detecting range, the baseline distance (T) and focal length (f) need to meet the detecting requirements.

Equation (4-2) can be transformed into

$$Z = \frac{fT}{d} \tag{7-1}$$

where $d = x_l - x_r$, then

$$\frac{\Delta Z}{\Delta d} = \frac{-fT}{d^2} \Rightarrow |\Delta Z| = \frac{fT\Delta d}{d^2} \quad \textbf{(7-2)}$$

replace d with $d = \frac{fT}{Z}$, then

$$|\Delta Z| = \frac{Z^2 \Delta d}{fT} \quad \textbf{(7-3)}$$

For the proposed system, the angle of view of lens is 24°x18°, the resolution of detector is 640x480pixels, the size of detector is 8mmx6mm and the baseline distance of two cameras is about 0.84m. Then the relationship between disparity, detecting range and distance resolution are calculated and shown in Figure 7-3.

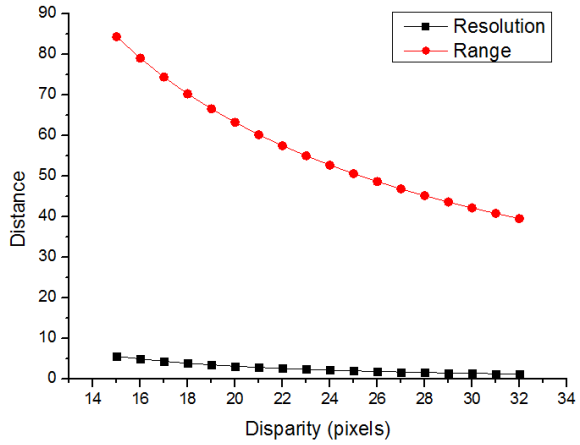


Figure 7-3: Depth and Disparity

The axis X is the disparity while axis Y is the distance. The detection distance has been defined in the system specification in Chapter 3.1, which is 40~80m. With a minimum disparity of 16 pixels, the detection range can reach 80m and the distance resolution at that distance is about 5 m. The distance resolution can reach 1.2 m when the detection range is 40m.

7.2 Stereo Calibration

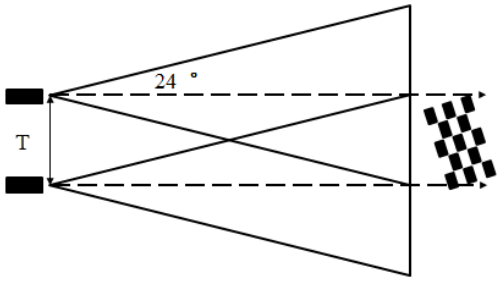
7.2.1 Chessboard

As discussed in Chapter 2.3, several calibration patterns can be used to calibrate the thermal infrared camera. In order to use the function in OpenCV, a printed chessboard is adopted as the calibration pattern.

The size of the chessboard needs to be figured out first. The baseline distance and field of view will influence the common area of two cameras at a certain distance, which are shown in Figure 7-4 (a). For a small chessboard, the chessboard can only occupy a few pixels in the image, which will influence the precise of calibration. Thus, the blocks in the chessboard should be big enough

to be detected by the algorithm. In order to occupy half of the image, the length of the chessboard need to be at least T m (T is the baseline distance). In this research, a large printed chessboard is made for the calibration, shown in Figure 7-4 (b). The length of a side of each block is 9 cm.

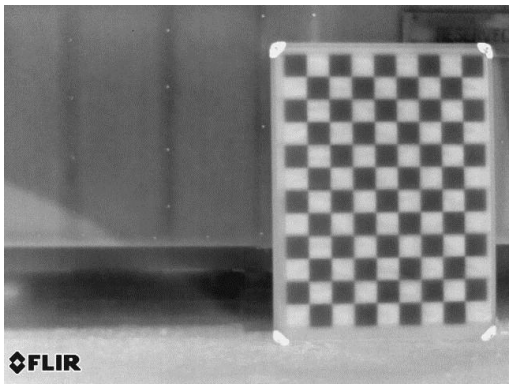
The chessboard needs to be heated evenly before calibration. Sun light is an ideal radiation source for heating the chessboard. Figure 7-5 shows the chessboard captured by the left and right infrared camera after heating by the sun light. Since the white and black colour have different absorption rate of heat, the slight temperature difference in the chessboard can be detected by the infrared cameras.



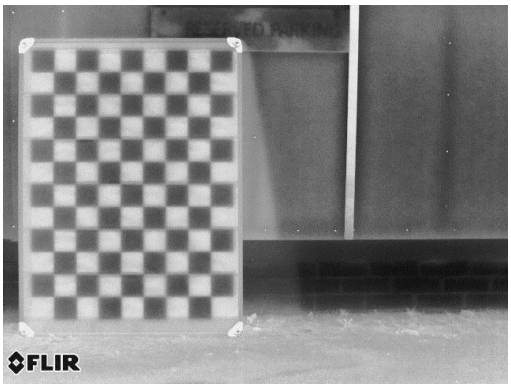
(a) Chessboard size according to baseline distance

(b) A large chessboard

Figure 7-4: The Common Field of Two Cameras



(a) Thermal image captured in the left camera



(b) Thermal image captured in the right camera

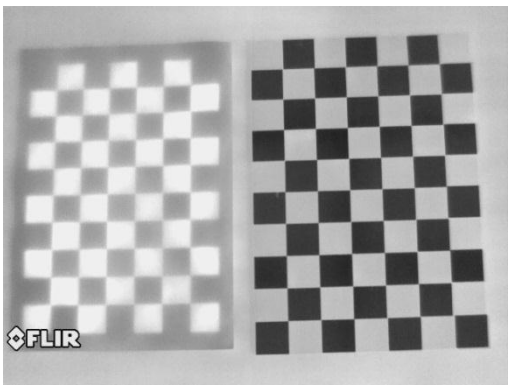
Figure 7-5: Thermal Image of Heated Chessboard

Meanwhile, a chessboard made of two materials was made in this research. It aims to improve the contrast of thermal images. The chessboard was made of plastic and aluminium pieces. Due to the different infrared reflection rate of different materials, the contrasts of white and black blocks were enhanced in the infrared image.

A printed chessboard and a chessboard made of aluminium and plastic pieces are shown in Figure 7-6 (a). The thermal images of the two chessboards are

shown in Figure 7-6 (b). Both chessboards were heated by the sunlight and a thermal image was taken for the two chessboards. It can be seen that the chessboard made of plastic and aluminium provided a sharper edge than the printed chessboard. However, since the plastic and aluminium chessboard was not big enough, a printed large chessboard was used in the calibration.

(a) Printed chessboard (left) and chessboard made of plastic and aluminium pieces (right)



(b) Thermal image of printed chessboard (left) and chessboard made of plastic and aluminium pieces (right)

Figure 7-6: Two kinds of Chessboard

7.2.2 Calibration Result

Figure 7-7 shows the calibration process of the system. Thermal images of chessboard were taken out-door. The big chessboard was laid against the wall. After being heated by the sunlight for a while, a series of images were taken from different angles. Then, the captured images were processed on to the laptop. Figure 7-8 shows images of the chessboard with corners found and lined in the process of calibration. The camera intrinsic matrices, lens distortion coefficients, rotation matrix, translation vector, essential matrix and fundamental matrix were obtained after calibration. Table 7-2 and Table 7-3 show these parameters obtained in the calibration.

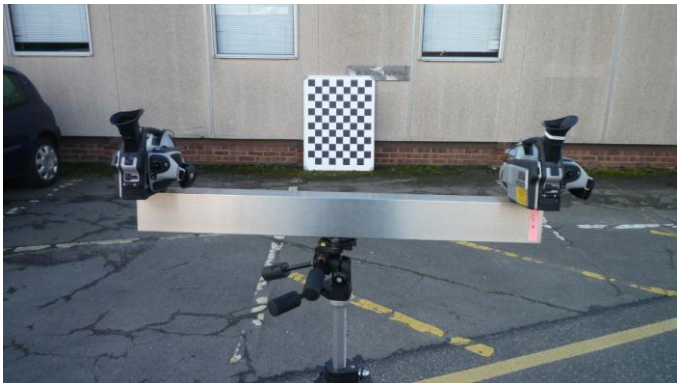


Figure 7-7: Thermal Infrared Camera Calibration

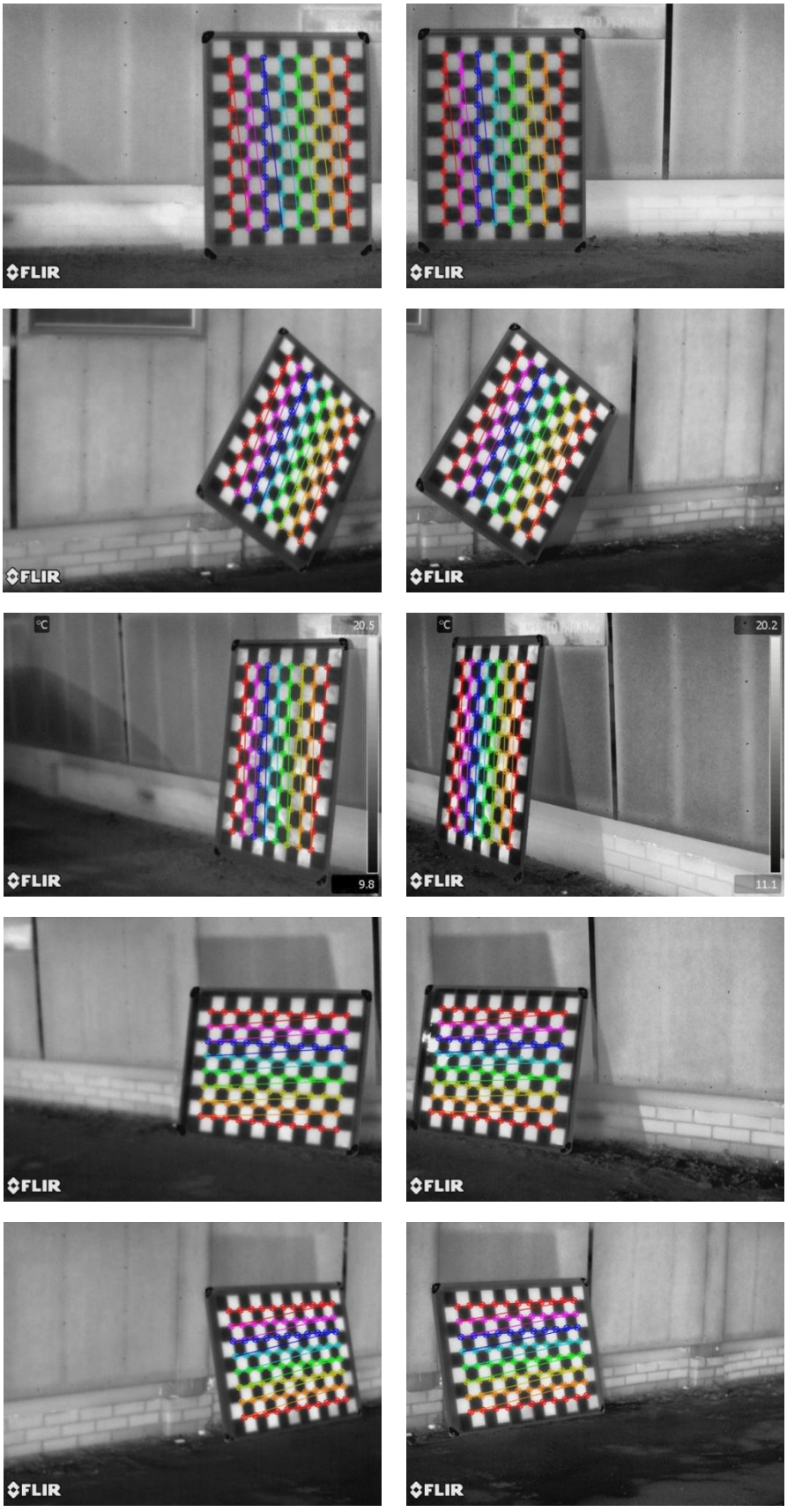


Figure 7-8: Thermal Images Captured in Stereo Calibration

Table 7-2: Intrinsic Calibration Results

Calibration Parameters		Left camera	Right camera
Focal length (pixels)	f_x	1531	1531
	f_y	1532	1532
Principal point (pixels)	C_x	330.14	331.32
	C_y	239.21	239.78
Lens distortion coefficient	k_1	-0.034	-0.17
	k_2	-12.07	-4.35
	k_3	0	0

Table 7-3: Relative Extrinsic Calibration Results

Rotation Matrix	**R**	$\begin{bmatrix} 0.99 & -0.0078 & -0.038 \\ 0.0077 & 0.99 & -0.0035 \\ 0.038 & 0.0032 & 0.99 \end{bmatrix}$
Translation Vector	**T**	$\begin{bmatrix} 83 \\ 1.47 \\ 0.32 \end{bmatrix}$
Essential Matrix	**E**	$\begin{bmatrix} -0.055 & 0.31 & -1.47 \\ 2.90 & 0.27 & 83.02 \\ 0.83 & -83.08 & 0.23 \end{bmatrix}$
Fundamental Matrix	**F**	$\begin{bmatrix} -5.01e-8 & 2.89e-7 & -2.11e-3 \\ 2.64e-6 & 2.47e-7 & 1.14e-1 \\ 5.49e-4 & -1.15e-1 & 9.99e-1 \end{bmatrix}$

7.3 Testing Result

7.3.1 Distance Estimation Error

The distance estimation error was evaluated first before testing at the airport. A person was regarded as an obstacle standing at different positions, shown in Figure 7-9. The distances of these positions were measured by the steel tape-measure. The range of these positions is 10m~60m. After capturing images of

the people at these positions, the distances of people could be calculated by the system. Figure 7-10 shows the measured distance and the actual distance. The system's detecting error and theoretical detecting error are compared in Figure 7-11. The theoretical error is calculated by equation (7-3).

Experiments shows that generally the detecting error line is less than the theoretical error line. However, the maximum error of detecting error is about 2 meter at the distance of 40 metres, larger than the theoretical error.

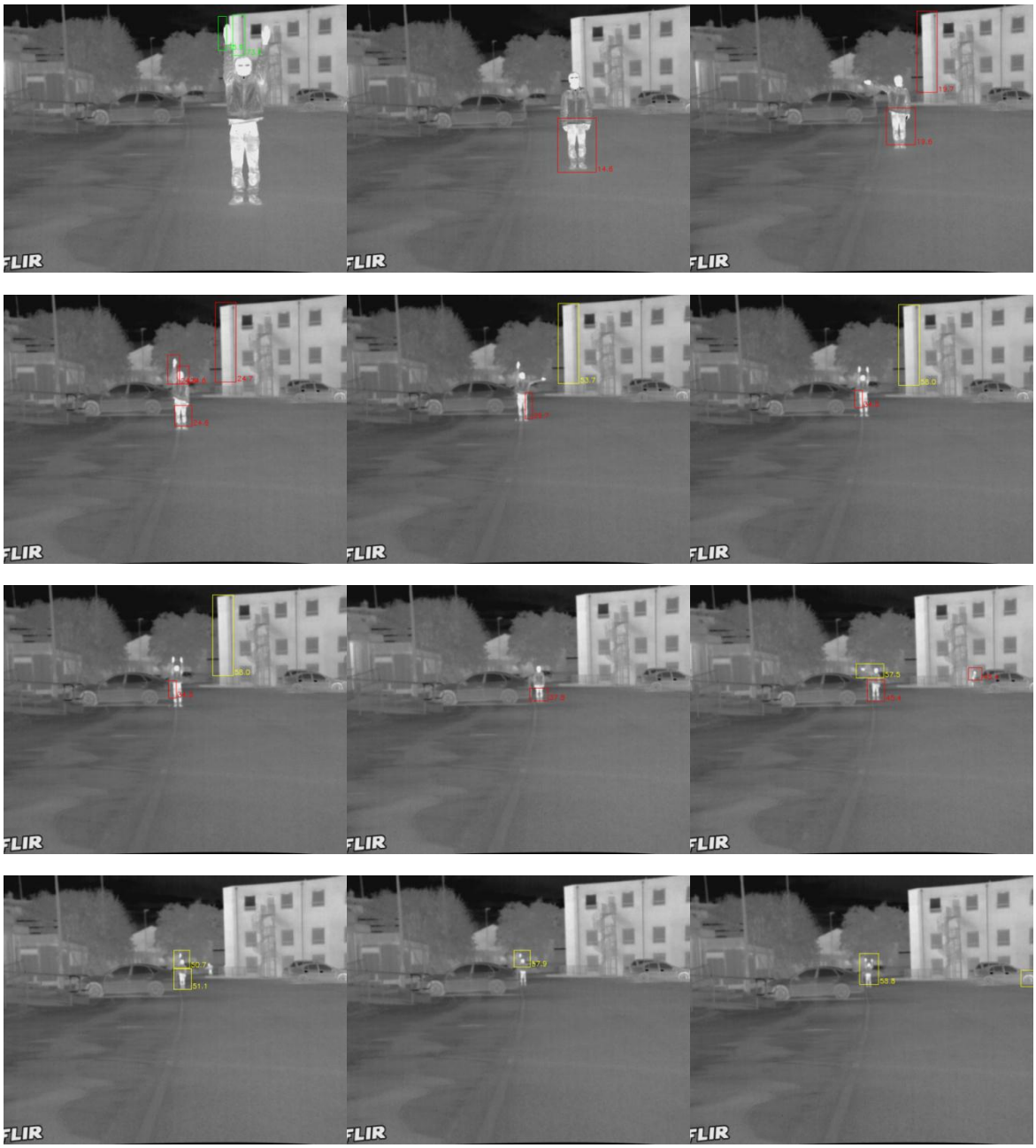


Figure 7-9: Distance Estimation Error Testing

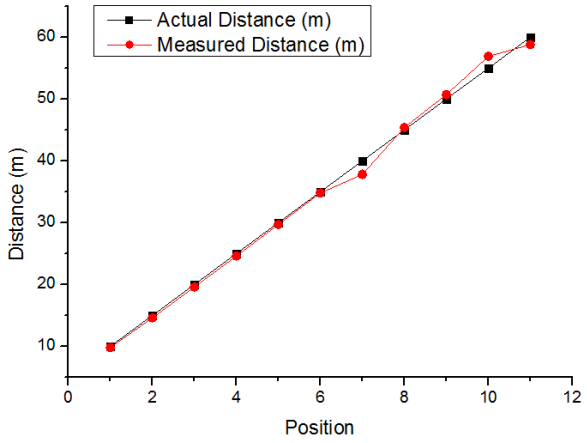


Figure 7-10: Actual Distance and Measured Distance

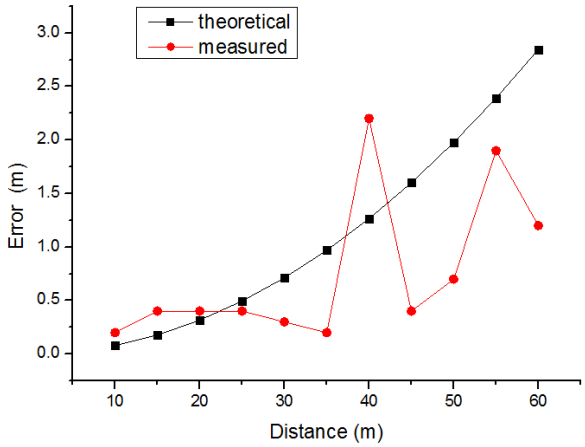


Figure 7-11: Error of Distance Estimation

7.3.2 Imaging Under all Weathers

Thermal images were taken under all kinds of weather conditions in order to evaluate the weather influence to the system performance. These weather conditions included fog, night, rain and sunshine. The results are shown in Figure 7-12~Figure 7-16. Images in the left column were captured by the optical camera and images in the right column were captured by the thermal infrared camera.

Figure 7-12 shows the people in fog. It is hard to identify the person in the optical image. However, the image of people is more obvious in the thermal image. The building is blurred in the thermal images as the fog absorbed the IR rays emitted from the building.

Figure 7-13 shows the airplanes in fog. Airplanes are recognizable in the optical image while nothing is recognizable in the thermal image. This is because airplanes parked on the ground were cooled down to the temperature of environment.

Figure 7-14 shows the objects at night. It can be seen that the even the illumination is poor, the thermal image provided a clearer outline of the objects than the optical image. Also, similar to Figure 7-13, when aircrafts are parked at the airport and have cooled down to the environment temperature, it is difficult for a thermal infrared camera to detect them.

Figure 7-15 shows the twin propeller in rain. In that condition, the thermal image is more blurred than the optical image due to the absorption of infrared rays. The outline of aircraft is not obvious in the thermal image. However, the hot point (two engines) is still recognizable.

Figure 7-12: People in Fog

Figure 7-13: Airplanes in Fog

Figure 7-14: Airport at Night

Figure 7-15: Airport in Rain Day

Figure 7-16: Airport on Sunny Day

Figure 7-16 shows the airport on sunny day. The pictures show that this condition is ideal for the optical camera. Since the illumination is adequate, the camera can provide clear images. However, the objects and environment in the thermal image is blended due to the reflection of sunlight. It can be seen that the sunlight reflected by the aircraft body, ground and cloud make them almost have the same measured temperature, result in deterioration of image quality, making background elimination difficult. The cloud or ground may still remain in the thermal image after background elimination.

7.3.3 Temperature Range

Human body's temperature was measured in different positions in cloudy and fog. This test was carried out in order to compare the temperature drop measured under different weather conditions. The data measured can be used to optimize the setting of membership function in the obstacle classification, which is described in the chapter 6.3. It allows adjustment of membership function of a certain obstacle in different weather conditions.

The first test was carried out in one afternoon, the weather was cloudy. Figure 7-17 shows the attenuation curve of human body measured. Two curves represent the temperature measured from the left and right camera. It can be seen the temperature decreased steadily between 10 metres and 100 metres.

The temperature is about 30℃ at the distance of 10 metres and 24℃ at the distance of 100 metres, dropped about 6℃.

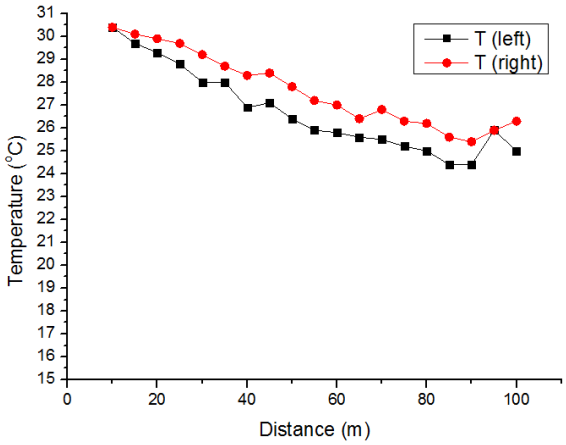


Figure 7-17: Human body's Temperature (Cloudy)

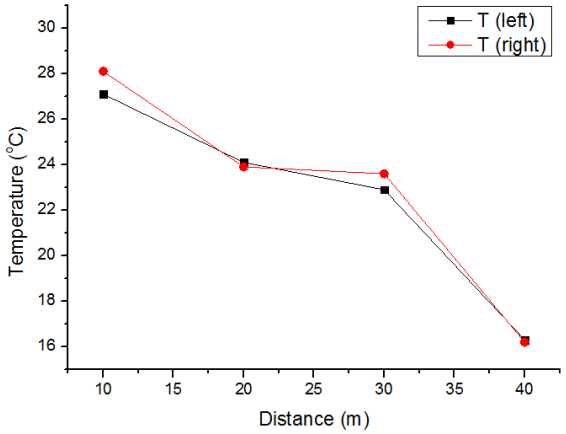


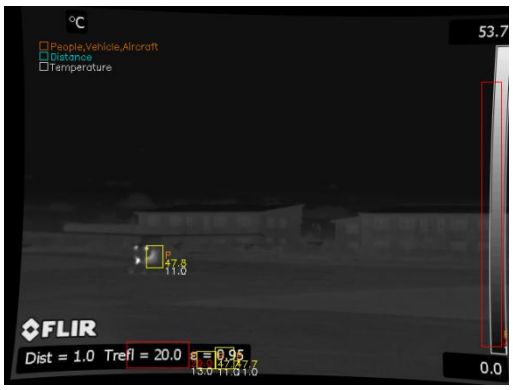
Figure 7-18: Human body's Temperature (Fog)

Figure 7-18 shows the attenuation curve of human body temperature measured in fog. It can be seen that the measured temperature decreased significantly when the distance increases. At the distance of 40 metres, the temperature dropped about 12℃. The environment measured by the infrared camera tis about 12~13℃. It can be seen the fog has a high absorption rate of infrared ray.

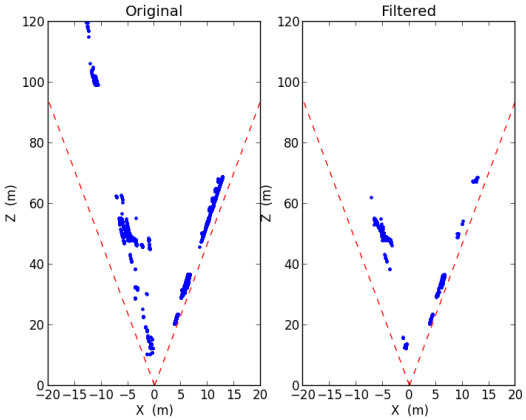
Thus, the parameters of membership function can be set according to different weathers.

7.3.4 Testing at the Airport

The final testing was carried out at the Cranfield Airport. The system was placed beside the taxi-way. Aircrafts and vehicles passing by were captured and stored by the two cameras. Then, the images or videos were processed offline to estimate the distance of obstacle.

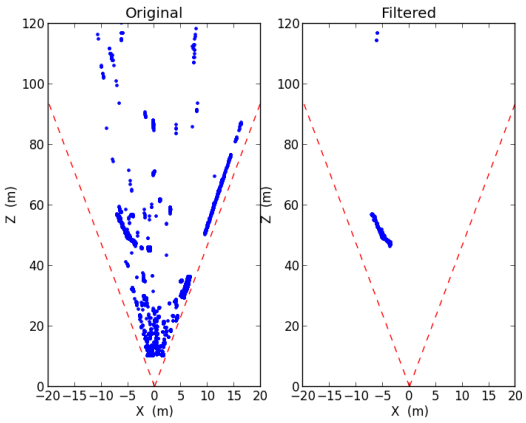


(a) Airplane



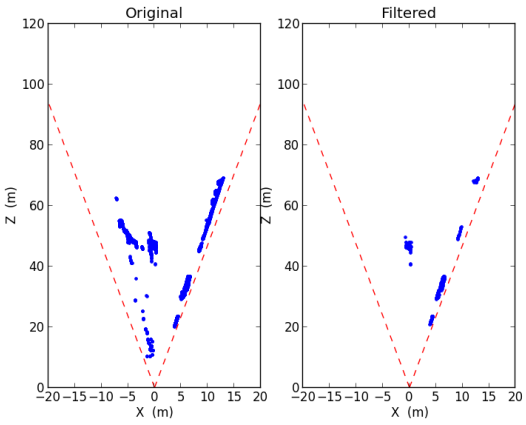
(b) Depth map according to Figure (a)

(c) Vehicle



(d) Depth map according to Figure (c)

(e) People



(f) Depth map according to Figure (e)

Figure 7-19: Obstacle Detecting and Depth Map Filtering

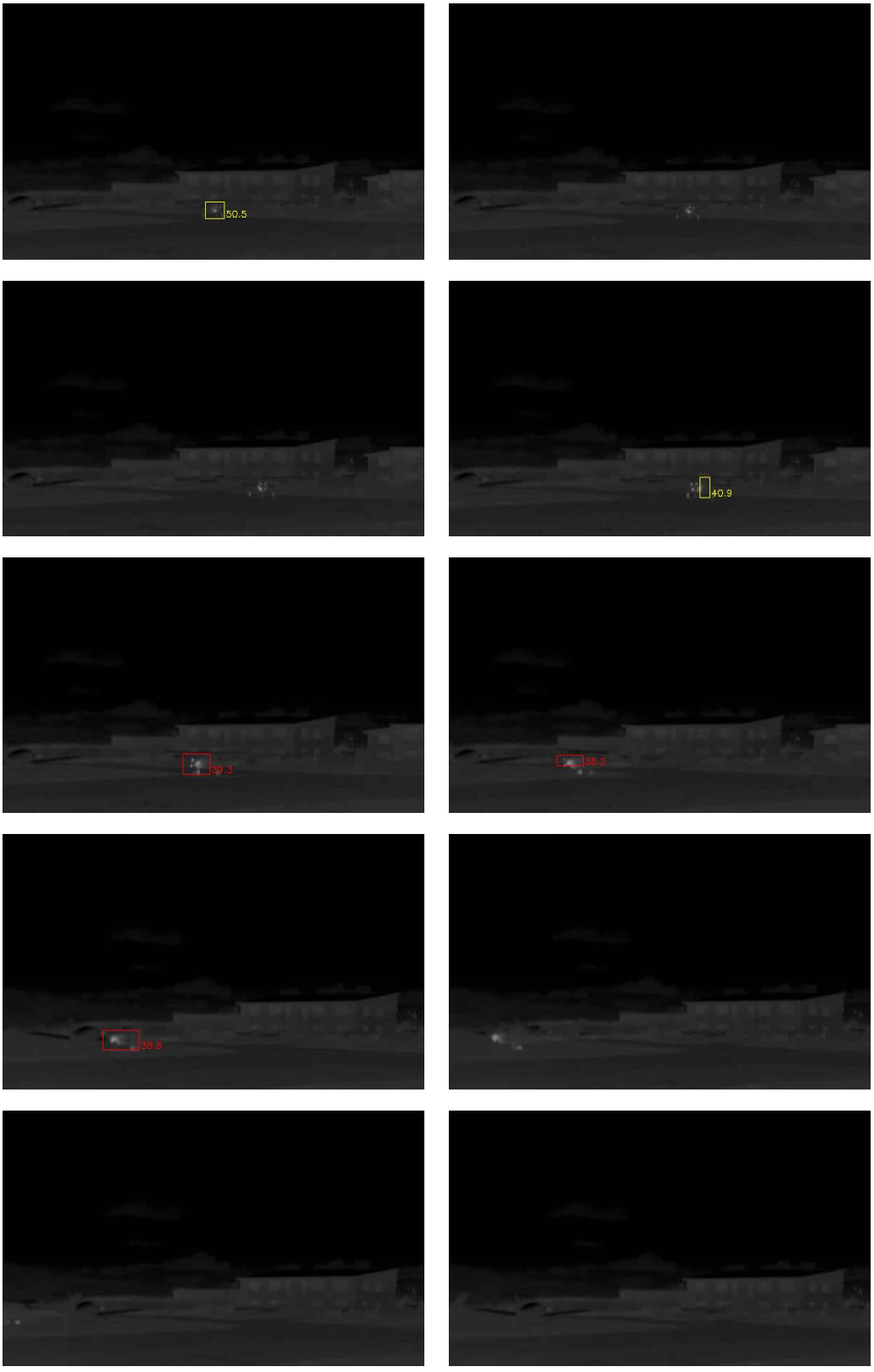


Figure 7-20: Moving Target Detecting

Figure 7-19 shows the detection result of airplanes, truck and people captured during the tests. The left column shows the rectified thermal image captured by the left camera of the system. The right column shows the depth map of obstacle viewed in XZ plane. Both original depth map and filtered depth map are shown. Since the position information is not added to the algorithm in this test, the only factor that influences the result is temperature. It can be seen that the airplane was misclassified as people. When the matched points have low temperature, the aircraft is misclassified as other obstacle. Although most of the noise was filtered, there were still some points are left in the depth map, such as the grey scale bar and subtitle.

In Figure 7-20, a series images show the distance estimation of a moving aircraft. These image sequences were converted from videos recorded by two cameras. The purpose was to testify the obstacle detecting performance of a moving target. However, the synchronization was a big problem. Since two thermal infrared cameras are not designed for working together online, the synchronization is depended on manual operation, which is activated by pushing the record button simultaneously. Obviously, this kind of synchronization cannot assure that the two cameras capture the scene simultaneously. In fact, the result of the testing indicates large distance estimation error. The error became even larger when the aircraft or vehicle was moving. In several images, the obstacle was not detected by the algorithm.

7.3.5 Simulation of Obstacle Classification

The concept of adding position information into obstacle classification was simulated in MATLAB. A total number of 426 thermal images captured in the Cranfield Airport were used in the simulation, including 157 images of Aircraft, 144 images of vehicle and 125 images of people.

The position information was given by a pseudo GPS coordinate. Pseudo GPS positions of people, vehicle and aircraft were distributed randomly in apron, taxi-way and runway, which are shown in Figure 7-21. The distribution of GPS positions are restricted by the obstacle types. For instance, the positions of people are distributed in apron while positions of aircrafts can be apron, taxi-way and run-way.

People	Apron		
Vehicle	Apron	Taxi-way	
Aircraft	Apron	Taxi-way	Run-way

Figure 7-21: Position Distribution of Objects

Figure 7-22 shows the map modified from a map of Cranfield Airport. The map is converted into a grey level map. The area of the run-way, taxi-way and apron are distinguished by different grey levels. In the simulation, the value of grey is used to decide the area in which the system is.

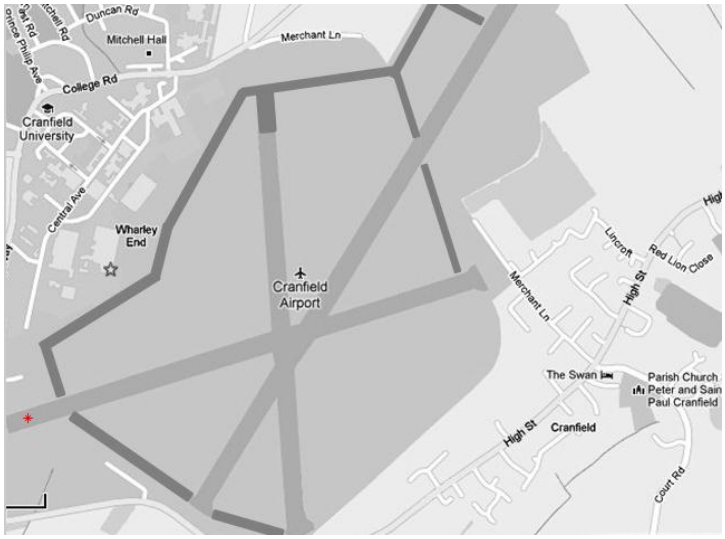


Figure 7-22: Map of Airport

The algorithm flow of classification is show in Figure 7-23. After received the pseudo GPS signal, the system will locate the area in the map through the received position information. Then, membership functions are selected according to the position information. Next, the type of obstacle is classified according to the obstacle's temperature. Since there is only one object of interest in each sample image, the highest temperature in the image is taken as the obstacle's temperature.

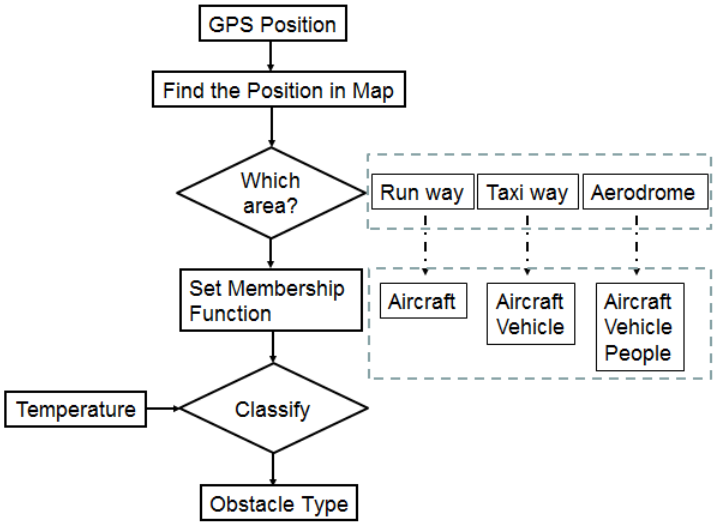


Figure 7-23: Concept of Obstacle Classification

Table 7-4 shows the result of the simulation. It can be seen that the overall accuracy of classification is enhanced 6% by involving the position information.

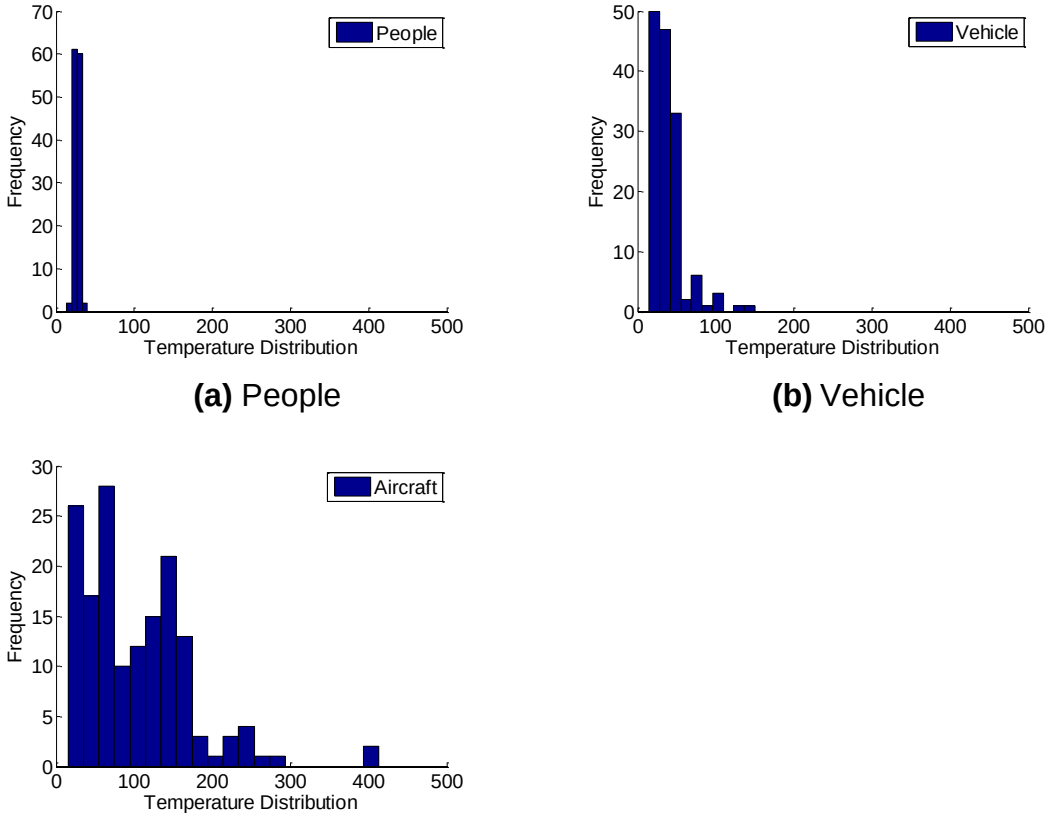
The classification accuracy of aircraft is low (about 54%). Two factors may contribute to the low accuracy. First, the temperature distribution of aircraft has a wide range, which can be seen in Figure 7-2. There are temperature overlaps between people, vehicles and aircrafts. The temperature overlaps may result in false classification. Second, since the aircraft could be in run-way, taxi-way or

apron, the position information has no effect in the classification for aircrafts. Thus, the combination of these two factors results in low accuracy. However, the images of aircrafts captured in the test are almost small aircrafts which has a relatively low temperature. For large passenger aircraft, the high temperature of jet engines will decrease the temperature overlaps.

It also can be seen the classification accuracy of vehicle is also low (55%) when no position information is involved. However, the classification accuracy of vehicle is enhanced about 23% by involving the position information. The position information can constrain the obstacle types which the accuracy of classification can be enhanced.

The classification accuracy of people remains same when position information is involved. Due to the concentrating distribution of temperature and relatively low temperature, the classification accuracy of people is relatively high (90%).

Therefore, the overall classification accuracy can be improved by involving the position information. For lager passenger aircrafts, the classification accuracy of aircrafts can also be improved.



(a) People

(b) Vehicle

(c) Aircraft

Figure 7-24: Temperature Distribution of Objects

Table 7-4: Performance Comparisons

	Accuracy (Fuzzy logic Only)	Accuracy (Position Involved)
People	90%	90%
Vehicle	55%	78%
Aircraft	54%	54%
Total	65%	71%

8 DISCUSSION AND CONCLUSION

8.1 Strengths and Limitations

Obstacles emitting infrared rays can be detected by the proposed system in poor weather conditions. The thermal infrared camera has a better image quality than the optical camera in low illumination conditions such as in fog and at night. However, the performance of thermal infrared cameras deteriorates when there is strong reflection of sunlight. Due to the reflection of sunlight from the ground and cloud, the measurement of temperature of ground or cloud was much higher, which is negative for background elimination.

The obstacle classification is realized by employing fuzzy logic and position information. Membership functions are built according to the temperature distribution of obstacles. By converting temperature into probability, it is easy to build general temperature models for obstacles. When the position information (GPS signal) is involved in the classification, the membership functions are constrained by the types of obstacles in that area. For instance, the measured temperature of a small aircraft may have the same temperature as a vehicle. When the position information indicates the system is on a runway, then only membership function of aircraft is selected. Thus, the obstacle will be classified as an aircraft.

Obstacles are highlighted by bounding box. Most of the noise on the depth map can be filtered based on contour area. In the corresponding process, mismatched points can be the noise in the image. The contour area based filtering is to filter small contours and large contours that cannot be obstacles. The testing shows an acceptable result of filtering.

Since the two cameras are designed mainly for capturing image individually, the system is not an on-line system and the synchronization of the two cameras is an issue. The synchronization can be achieved by pressing the shutters in both cameras simultaneously. Obviously, it is difficult to ensure the two cameras capture the image at the same time, especially when the target is moving. In the experiment, the first attempt was to detect the distance of moving target. Videos were shot and converted to images for processing. However, the result turned out to be unacceptable due to the synchronization issue.

The code developed in Python shows a fast image processing speed. The average image processing time is about 0.4s per frame, much faster than the image processing speed in [19]. However, for the application of real-time processing, the processing speed should be less than 1/(frame rate) s.

8.2 Contributions

As a result of this research, the main contributions to the field can be identified:

1. An obstacle detection system using thermal imaging sensor is developed for large commercial aircraft. In [41], the stereo vision system is proposed to be used in the obstacle detection for large commercial aircraft and it is suggested

that future research needs to be done using thermal infrared cameras. In this thesis, thermal infrared cameras are employed in the stereo vision system to detect obstacles. Moreover, according to feature of thermal imaging, proper image processing algorithms are selected and implemented. The image processing time is reduced by programming in a compiled language (Python).

2. An obstacle classification method based on fuzzy logic is proposed in this thesis. Membership functions are chosen based on the distribution of temperature measured. It turns the temperature range into probability between 0 and 1 in the classification. When taking position information into consideration in the classification, selecting membership function is easier than defining the temperature range of obstacles.

3. Contour area based filtering in the depth map is proposed in this project. For large contour area in the depth map, the actual area of obstacle is calculated roughly by projecting the contour area into the obstacle panel (the focus length and obstacle distance are calculated in the stereo vision technology). By setting a threshold of contour area, most of the noises can be filtered.

8.3 Suggestions of Future Work

Due to time limitation it was not possible to fully accomplish the work needed in order to fully develop the system in this project, consequently suggestions for future work in order to complete and enhance the project is as follows:

Due to the limitation of thermal infrared cameras borrowed to be used in this project, two cameras were synchronized manually instead of computer and images were processed offline. Thus, the inaccurate synchronization of two thermal infrared cameras may cause significant distance estimation error. Also, the system developed is not an on-line and real time system. In order to make the system work online, proper thermal infrared cameras which can capture the scene simultaneously under the control of sync clock need to be employed. In order to make the system operate in real-time, more work needs to be done to reduce the image processing time. Although image processing time has been reduced to 0.4s per frame in this project, it is still not fast enough. Reducing image processing time can be achieved by optimizing algorithms and using more advanced computer hardware.

Since the strong reflection of sunlight has a negative effect on thermal image, combination of stereo vision system using a pair of optical camera and a pair of thermal infrared camera would produce a better performance and reliability in all weather conditions. The infrared camera and optical camera have complementary properties. An optical camera is suitable for weather when illumination is sufficient while a thermal infrared camera perform well in poor weather conditions. Thus, in the future work, the data fusion algorithm needs to be developed in the future work.

The obstacle classification based on temperature and GPS position information was simulated in this project. Real implementation can be done in the future work, in which a GPS receiver needs to be integrated in the system. Also, maps

of airports are also needed in order to be able to identify different areas of the airports.

8.4 Conclusion

The main objectives of this research have been met. The system performance specification has been defined. Image Processing algorithms for infrared image were chosen and implemented. Distance of obstacle is estimated by the stereo vision technology. Obstacle can be detected and classified according to their temperature. Moreover, the image processing time is reduced by programming in Python. The whole system was tested at the airport and the result shows the system is capable of detecting static and low speed obstacles under poor weather conditions. Also, a concept of obstacle classification involving position information is proposed and simulated. If implemented in an aircraft, the pilots' situational awareness can be enhanced by the proposed system. Finally, further work is needed in order to develop the system into an on-line and real time system.

REFERENCES

[1] "BBC NEWS | UK | Why fog causes airport chaos." [Online]. Available: http://news.bbc.co.uk/1/hi/uk/6200527.stm. [Accessed: 27-Nov-2012].

[2] "Google Maps." [Online]. Available: https://maps.google.co.uk/maps?hl=en&tab=ll. [Accessed: 14-Dec-2012].

[3] "Surface Movement - Air Traffic Control Systems - Easat Antennas." [Online]. Available: http://www.easat.com/en/systems/air-traffic-control/surface-movement-radar. [Accessed: 21-Nov-2012].

[4] "File:Ground radar EHAM.jpg - Wikipedia, the free encyclopedia." [Online]. Available: http://en.wikipedia.org/wiki/File:Ground_radar_EHAM.jpg. [Accessed: 21-Nov-2012].

[5] "Communications provide direct links between all equipped users and ATC. - Image - Airport Technology." [Online]. Available: http://www.airport-technology.com/features/feature619/feature619-2.html. [Accessed: 21-Nov-2012].

[6] "SKYbrary - A-SMGCS." [Online]. Available: http://www.skybrary.aero/index.php/A-SMGCS. [Accessed: 21-Nov-2012].

[7] F. Esposito, G. Rufino, A. Moccia, and I. Introduction, "Real-Time Detection of Fire Hotspots from Mini-UAV Based, Thermal InfraRed / VIS-NIR Hyperspectral Image Data," no. April, pp. 1–18, 2009.

[8] L. Matthies and a. Rankin, "Negative obstacle detection by thermal signature," *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453)*, vol. 1, pp. 906–913, 2003.

[9] A. Treptow, G. Cielniak, and T. Duckett, "Real-time people tracking for mobile robots using thermal vision," *Robotics and Autonomous Systems*, vol. 54, no. 9, pp. 729–739, Sep. 2006.

[10] I. R. C. Performance, "Seeing through fog and rain with a thermal imaging camera."

[11] T. Kato, Y. Ninomiya, and I. Masaki, "An obstacle detection method by fusion of radar and motion stereo," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 3, pp. 182–188, Sep. 2002.

[12] S. Sugimoto, H. Tateda, H. Takahashi, and M. Okutomi, "Obstacle detection using millimeter-wave radar and its visualization on image sequence," *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, pp. 342–345 Vol.3, 2004.

[13] A. Leykin, “Thermal-Visible Video Fusion for Moving Target Tracking and Pedestrian Classification,” *Computing*, 2007.

[14] N. Yonemoto, K. Yamamoto, and K. Yamada, “Performance of obstacle detection and collision warning system for civil helicopters,” no. September 2004.

[15] M. Perrollaz, C. Roy, N. Hauti, and D. Aubert, “Long Range Obstacle Detection Using Laser Scanner and Stereovision,” pp. 182–187, 2006.

[16] H. Nakai, N. Takeda, H. Hattori, Y. Okamoto, and K. Onoguchi, “A practical stereo scheme for obstacle detection in automotive use,” *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, no. 1, pp. 346–350 Vol.3, 2004.

[17] a. Bensrhair, a. Bertozzi, a. Broggi, a. Fascioli, S. Mousset, and G. Toulminet, “Stereo vision-based feature extraction for vehicle detection,” *Intelligent Vehicle Symposium, 2002. IEEE*, vol. 2, pp. 465–470.

[18] P. Trisiripisal, M. R. Parks, and A. L. Abbott, “Stereo Analysis for Vision-based Guidance and Control of Aircraft Landing,” no. January, pp. 1–17, 2006.

[19] J. Gauci, “Obstacle Detection around Aircraft on Ramps and Taxiways through the use of Computer Vision,” no. August, pp. 1–23, 2009.

[20] A. Rankin, A. Huertas, L. Matthies, M. Bajracharya, C. Assad, S. Brennan, P. Bellutta, and G. Sherwin, “Unmanned ground vehicle perception using thermal infrared cameras,” 2011.

[21] T. A. Williamson, “A High-Performance Stereo Vision System for Obstacle Detection,” 1998.

[22] B. Hulin and S. Schüßler, “Concepts for Day-Night Stereo Obstacle Detection in the Pantograph Gauge,” vol. 49, no. 0, pp. 449–454, 2007.

[23] B. A. Barsky, D. R. Horn, S. A. Klein, and J. A. Pang, “Camera Models and Optical Systems Used in Computer Graphics : Part I , Object-Based Techniques,” pp. 246–255.

[24] C. Kolb, D. Mitchell, and P. Hanrahan, “A realistic camera model for computer graphics,” *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques - SIGGRAPH ’95*, pp. 317–324, 1995.

[25] S. Vidas, R. Lakemond, S. Denman, C. Fookes, S. Sridharan, S. Member, and T. Wark, “A Mask-Based Approach for the Geometric Calibration of Thermal-Infrared Cameras,” vol. 61, no. 6, pp. 1625–1635, 2012.

[26] S. Yahyanejad and J. Misiorny, “Lens distortion correction for thermal cameras to improve aerial imaging with small-scale UAVs,” *Robotic and Sensors*, 2011.

[27] R. Yang and Y. Chen, “Design of a 3-D Infrared Imaging System Using Structured Light,” *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 2, pp. 608–617, Feb. 2011.

[28] G. of Canada, T. Canada, and C. A. Safety and Security Group, “Taxi Check and Procedures,” Jan. 2004.

[29] R. V. Levine and a. Norenzayan, “The Pace of Life in 31 Countries,” *Journal of Cross-Cultural Psychology*, vol. 30, no. 2, pp. 178–205, Mar. 1999.

[30] B. J. Mohler, W. B. Thompson, S. H. Creem-Regehr, H. L. Pick, and W. H. Warren, “Visual flow influences gait transition speed and preferred walking speed.,” *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, vol. 181, no. 2, pp. 221–8, Aug. 2007.

[31] R. Jordan, M. A. Ishutkina, T. G. Reynolds, and W. Street, “A Statistical Learning Approach to the Modeling of Aircraft Taxi-Time,” vol. 298, no. 0704, 2010.

[32] S. Verma, S. Lozito, T. Kozon, D. Ballinger, and H. Resnick, “Procedures for Off-Nominal Cases : Very Closely Spaced Parallel Runway Operations,” 2008.

[33] “Airbus A380-800 Brake test.” [Online]. Available: http://www.youtube.com/watch?v=m1dv_y_3EK0. [Accessed: 29-Nov-2012].

[34] D. Industry, “The Ultimate Infrared Handbook for R & D Professionals.”

[35] “An Inside Look at Boeing’s 787 Dreamliner Flight Simulator | Airline Reporter | Blogging on the airline business.” [Online]. Available: http://www.airlinereporter.com/2012/03/an-inside-look-at-boeings-787-dreamliner-flight-simulator/. [Accessed: 16-Dec-2012].

[36] G. Bradski and A. Kaehler, *Learning OpenCV*. .

[37] Z. Zhang, “A Flexible New Technique for Camera Calibration,” vol. 1998, 2002.

[38] D.C.Brown, “Close Range Camera Calibration.pdf.” Photogrammetric Engineering, pp. 855–866, 1971.

[39] D. Scharstein, R. Szeliski, and R. Zabih, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” *Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV 2001)*, no. 1, pp. 131–140.

[40] R. L. Issue, “User ’ s manual ThermaCAM ™ B640 ThermaCAM ™ P640 ThermaCAM ™ SC640.”

[41] J. Gauci, “Obstacle Detection in Aerodrome Areas Through the Use of Computer Vision.”

[42] I. Nedeljkovic, C. Vi, and W. G. Vi, “IMAGE CLASSIFICATION BASED ON FUZZY LOGIC,” pp. 1–6.

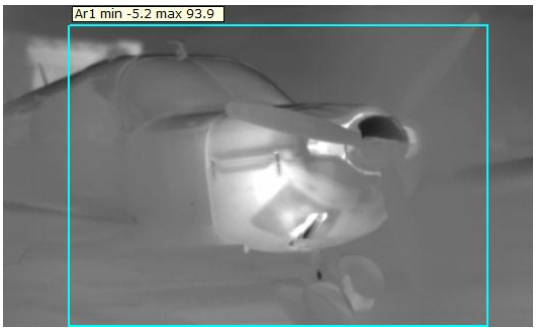
[43] M. Sharma, R. Gupta, D. Kumar, and R. Kapoor, “Efficacious approach for satellite image classification,” vol. 3, no. October, pp. 143–150, 2011.

[44] T. Çelik, H. Özkaramanlæ, and H. Demirel, “FIRE PIXEL CLASSIFICATION USING FUZZY LOGIC AND STATISTICAL COLOR MODEL,” pp. 1205–1208, 2007.

[45] W. K. Wong, Z. Y. Chew, C. K. Loo, and W. S. Lim, “An Effective Trespasser Detection System Using Thermal Camera,” *2010 Second International Conference on Computer Research and Development*, pp. 702–706, 2010.

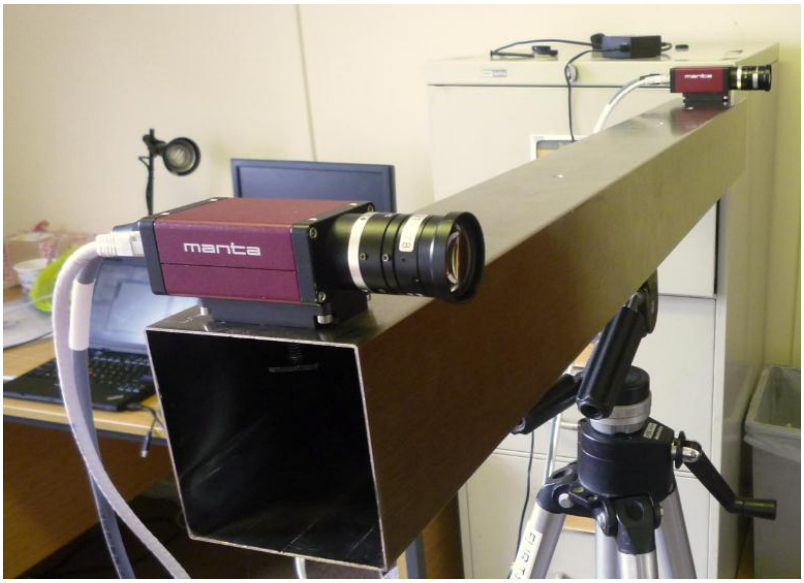
APPENDICES

Appendix A **Temperature of Objects**

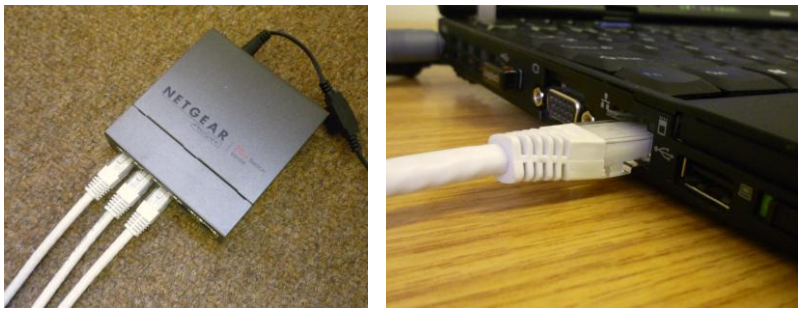


Appendix B **Optical Stereo Vision System**

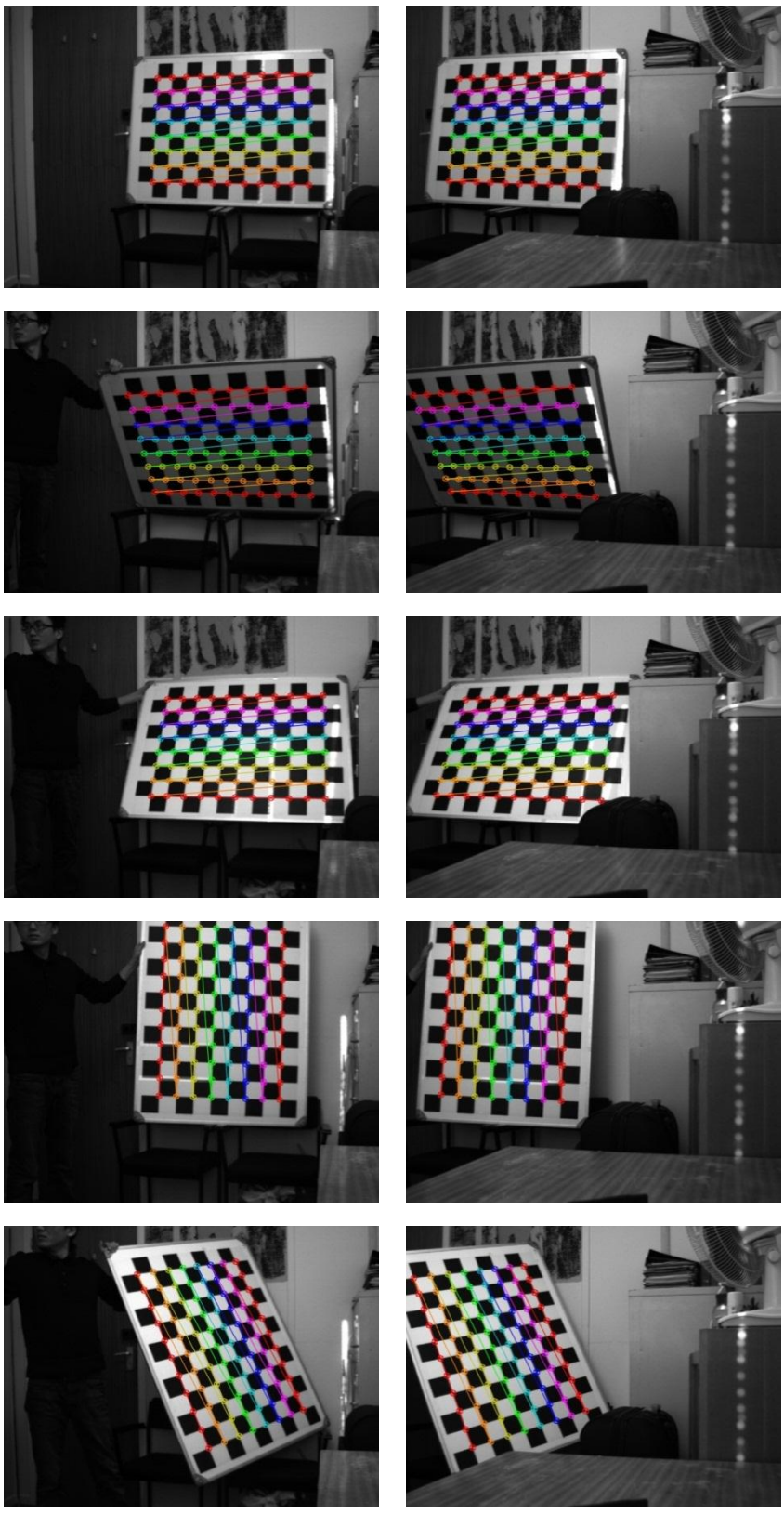
Since the loan of thermal infrared cameras is only one month (thermal infrared cameras are not available until October 2012), the stereo vision system was developed based on optical cameras first. Figure (a) below shows the optical stereo vision system developed in this project. Two cameras are mounted on the frame. The communication between computer and two cameras is through net cable. A router is employed to connect cameras and the computer, which is shown in Figure (b). Images captured in the system calibration are shown in Figure (d). Since the net cable can provide fast data transmitting and the two optical cameras are designed for online capturing, the system is capable of working in real-time. Due to the difference of image processing algorithms between optical image and thermal image, the obstacle detection and classification algorithms were not implemented in this system,.



(a) Stereo vision system based on optical cameras



(b) The link between camera and computer



(c) Optical Images Captured in the Stereo Calibration